

W P S R C H (TM)

Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by IntelliGenetics, Inc.

Msrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Tue Mar 19 09:53:17 1996; Maspar time 7.39 Seconds

555.014 Million cell updates/sec

Tabular output not generated.

Title: >US-08-404-832-2
Description: (1-567) from US08404832.pep
Perfect Score: 4162
Sequence: 1 MESSKMDSPGALQTNPLK.....IKDDTIFIKVYDTSDDPP 567

Scoring table: PAM 150
Gap 11

Searched: 62355 seqs, 7230759 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database:

a-geneseq20
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11 12:part12

Statistics: Mean 36.298; Variance 168.378; scale 0.216

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
1	132	3.2	527	7 R38099	Schistosoma mansoni	8.41e-02
2	132	3.2	527	3 R4183	Product of clone IrV-	8.41e-02
3	128	3.1	235	1 P94260	41KD protein of T. co	1.63e-01
4	128	3.1	320	1 P94366	41KD protein of T. co	1.63e-01
5	130	3.1	885	12 R6930	AMML chromosome inv(1	1.17e-01
6	125	3.0	1284	1 P8187	Sequence encoded by a	2.66e-01
7	118	2.8	252	1 P94369	Fusion protein congt.	8.27e-01
8	116	2.8	334	10 R51227	Membrane antigen pept	1.14e+00
9	117	2.8	337	11 R63681	Merosin amino acids 1	9.71e-01

10	117	2.8	385	1 P90994	B. subtilis sacu51 po	9.71e-01
11	118	2.8	476	8 R43563	Hyaluronan receptor.	8.27e-01
12	114	2.7	462	1 R05766	Portion of peptide an	1.57e+00
13	111	2.7	769	1 P90954	Yeast topoisomerase I	2.52e+00
14	111	2.7	870	6 R30729	pl00 protein from hum	2.52e+00
15	107	2.6	475	3 R15148	Ro/SSA autoantigen.	4.70e+00
16	110	2.6	816	12 R66931	AMML chromosome inv(1	2.94e+00
17	108	2.6	3685	1 P90290	Human Duchenne muscul	4.02e+00
18	103	2.5	200	9 R47245	Fragment of 101 kD pr	8.71e+00
19	103	2.5	318	5 R26943	P. falciparum LSA N-te	8.71e+00
20	104	2.5	674	6 R34130	Truncated PVX replica	7.47e+00
21	104	2.5	707	6 R34131	Truncated PVX replica	7.47e+00
22	104	2.5	1456	6 R34129	PVX replicase.	7.47e+00
23	104	2.5	1618	5 R27205	Human nestin.	7.47e+00
24	104	2.5	1618	11 R60127	Human nestin protein	7.47e+00
25	98	2.4	183	7 R39482	Human apolipoprotein	1.86e+01
26	101	2.4	316	5 R26941	P. falciparum LSA-R-NR	1.86e+01
27	98	2.4	333	7 R39480	Human apolipoprotein	1.86e+01
28	98	2.4	337	7 R39482	Human apolipoprotein	1.86e+01
29	98	2.4	342	7 R39487	Human apolipoprotein	1.86e+01
30	98	2.4	342	7 R39496	Human apolipoprotein	1.86e+01
31	98	2.4	342	7 R39491	Human apolipoprotein	1.86e+01
32	98	2.4	346	7 R39493	Human apolipoprotein	1.86e+01
33	99	2.4	361	9 R31052	Sequence encoded by t	1.60e+01
34	98	2.4	363	7 R39479	Human apolipoprotein	1.86e+01
35	98	2.4	373	7 R39486	Human apolipoprotein	1.86e+01
36	98	2.4	377	7 R39443	Human apolipoprotein	1.86e+01
37	98	2.4	377	7 R45244	Human apolipoprotein	1.86e+01
38	98	2.4	377	7 R39502	Human apolipoprotein	1.86e+01
39	98	2.4	377	7 R39501	Human apolipoprotein	1.86e+01
40	98	2.4	377	7 R45243	Human apolipoprotein	1.86e+01
41	98	2.4	377	7 R45242	Human apolipoprotein	1.86e+01
42	101	2.4	493	5 R26944	P. falciparum LSA gene	1.18e+01
43	99	2.4	576	12 R66929	AMML chromosome inv(1	1.60e+01
44	99	2.4	955	11 R57365	K39 polypeptide of le	1.60e+01
45	99	2.4	2101	9 R47173	Sequence of the inner	1.60e+01

ALIGNMENTS

RESULT	1
ID	R38099 standard; Protein; 527 AA.
AC	R38099;
DT	26-OCT-1993 (first entry)
DE	Schistosoma mansoni epitope.
KW	Antibody; protein epitope; glycanic epitope; surface;
KW	Schistosoma; myosin; C. elegans.
OS	Schistosoma mansoni.
PN	US5219566-A.
PD	15-JUN-1993.
PF	30-SEP-1988; 252075.
PR	30-SEP-1988; US-252075.
PR	09-JUL-1991; US-725804.
PA	(U100) UNIT JOHNS HOPKINS.
PI	Strand M;
PI	WPI; 93-205255/25.
DR	N-PEDBI; Q43521.
PT	New cDNA encoding polypeptide of Schistosoma mansoni - includes
PT	protein epitope(s) present on Schistosoma surface, used in
PT	vaccines against bilharziasis

Hayes
404832

309 105 243

PS Claim 1; Fig 1B; 11pp; English.
CC The polypeptide binds to antibodies which recognise protein epitopes,
CC but not glycanic epitopes, expressed on the surface of live
CC schistosoma of S. mansoni. The protein epitopes are common to a
CC 200 kD and a 38 kD glycoprotein of S. mansoni. These glycoproteins
CC are immunologically cross-reactive with myosin heavy chains from other
CC species. However, anti-myosin antibodies directed against myosin mols.
CC of other species are not cross-reactive with the surface epitopes of
CC S. mansoni.
SQ Sequence 527 AA;

Query Match 3.2%, Score 132; DB 7; Length 527;
Best Local Similarity 24.4%; Pred. No. 8.41e-02;
Matches 29; Conservative 33; Mismatches 51; Indels 6; Gaps 4;

Db 255 dneirtlgemagdmglknkknleeeenrtgeaqaeedkvnhlkakleat 314
QY 270 EKVSLLONEVEKNKSIQSLHNOICFELIEIRKEMLRNNSKILHQRVIDSOAKL 329
Db 315 demeennlareqkigdvksekrlleg-dlkatq-etvd-dlervkrd---leeqrlrke 367
QY 330 KEIDKEIRPFRRNWEADSMKSSVESLQNRVTELESVDKSGGVANRTLLSQRSD 388

RESULT 2
ID R14183 standard; Protein; 527 AA.
AC R14183;
DT 03-JAN-1992 (first entry)
DE Product of clone Irv-5.
KW Epitope; bilharziasis; antibodies; schistosomula; vaccine;
KW Immunogen; anti-Irv; fusion.
OS Schistosoma mansoni.
FH Key Location/Qualifiers
FT Domain 121..572
FT /label= alpha helix
FT Modified-site 160
FT /label= N-glycosite
FT Modified-site 218
FT /label= N-glycosite
PN US5051254-A.
PD 24-SEP-1991.
PF 30-SEP-1988; 252075.
PR 30-SEP-1988; US-252075.
PA (UYJO) JOHNS HOPKINS UNIV.
PI Strand M;
PI MPI; 91-303003/41.
DR N-PSDB; Q13982.
PT Polypeptide(s) specific for epitope(s) of Schistosoma mansoni -
PT used as vaccines against bilharziasis.
PS Claim 5; Fig 1B; 10pp; English.
CC The clone was obt. by screening an adult worm cDNA library in
CC lambda-gt10 with anti-Irv sera raised against a subset of adult worm
CC antigens. The sequence is thought to at least partially encode a
CC schistosome myosin. It has 48% identity with the myosin heavy
CC chain of C. elegans. The DNA can be used to produce a fusion protein
CC with beta galactosidase sequences. The product can be used as an
CC immunogen to produce vaccines against bilharziasis.
SQ Sequence 527 AA;

Query Match 3.2%, Score 132; DB 3; Length 527;

Best Local Similarity 24.4%; Pred. No. 8.41e-02;
Matches 29; Conservative 33; Mismatches 51; Indels 6; Gaps 4;

Db 255 dneirtlgemagdmglknkknleeeenrtgeaqaeedkvnhlkakleat 314
QY 270 EKVSLLONEVEKNKSIQSLHNOICFELIEIRKEMLRNNSKILHQRVIDSOAKL 329
Db 315 demeennlareqkigdvksekrlleg-dlkatq-etvd-dlervkrd---leeqrlrke 367
QY 330 KEIDKEIRPFRRNWEADSMKSSVESLQNRVTELESVDKSGGVANRTLLSQRSD 388

RESULT 3
ID P94260 standard; protein; 235 AA.
AC P94260;
DT 13-JUN-1990 (first entry)
DE 41kD protein of T. colubriformis.
KW Parasitic nematode; 41 kD protein.
OS Trichostrongylus colubriformis.
PN WO8900163-A.
PD 12-JAN-1989.
PF 06-JUL-1988; A00239.
PR 07-JUL-1987; A0-002940.
PA (BIOT-) Biotech Austr Pt, (CSIR) Commonwealth Sci and Ind Res Orgn.
PI Cobon GS, Austen RA, O'Donnell JU, Frenkel MJ, Kennedy WK, Savin KW,
PI Wagland BM;
DR MPI; 89-039628/05.
DR N-PSDB; N91210.
PT Protein derived from parasitic nematode species - used to provide
PT protective immunity against nematode parasites of man and animal.s
PS Disclosure; 57pp; English.
CC Recombinant organisms contg. DNA encoding the protein can be used to
CC produce polypeptide which is capable of protecting guinea pigs
CC against parasitism from T. colubriformis.
CC See also P94366-P94369.
SQ Sequence 235 AA;

Query Match 3.1%, Score 128; DB 1; Length 235;
Best Local Similarity 17.1%; Pred. No. 1.63e-01;
Matches 18; Conservative 35; Mismatches 49; Indels 3; Gaps 2;

Db 7 mkiekdnaldradaaeekvrgiteklerveeirdtqkmmgtendldkagelaaatsq 66
QY 299 IEIERKEMLRN--ESKILHQRVIDSOAEKELKEIRPFRRNWEA--DSMKSSVES 355
Db 67 leekkyggaeeavaalnrmrlleeeleeraerlkiateklee 111
QY 356 LQNRVTELESVDKSGGVANRTLLSQRSDMLSVHDIRLAD 400

RESULT 4
ID P94366 standard; protein; 320 AA.
AC P94366;
DT 13-JUN-1990 (first entry)
DE 41kD protein of T. colubriformis.
KW Parasitic nematode; 41 kD protein.
OS Trichostrongylus colubriformis.
PN WO8900163-A.
PD 12-JAN-1989.
PF 06-JUL-1988; A00239.

PR 07-JUL-1987; AU-002940.
PA (BIOT-) Biotech Aust Pt. (CSIR) Commonwealth Sci and Ind Res Orgn.
PI Cobon GS, Autech RA, O'Donnell LJ, Frenkel MJ, Kennedy WPK, Savin KW,
PI Magland BW;
DR WPI; 89-039628/05.
DR N-PSDB; N91211.
PT Protein derived from parasitic nematode species - used to provide
PT protective immunity against nematode parasites of man and animals.
PS Claim 4; page 39; 57pp; English.
CC Recombinant organisms contg. DNA encoding the protein can be used to
CC produce polypeptide which is capable of protecting guinea pigs
CC against parasitism from T. colubriformis.
CC See also P94260, and P94368-P94369.
SQ Sequence 320 AA;

Query Match 3.1%; Score 128; DB 1; Length 320;
Best Local Similarity 17.1%; Pred. No. 1.63e-01;
Matches 18; Conservative 35; Mismatches 49; Indels 3; Gaps 2;

Db 11 mkiakdnaldadaeekvrtgkierveeeldrtqkkmqetndlkagedlaatsq 70
QY 299 IEIEROKEMLRNN--ESKIIHLQVIDSQAEKIKELDKREIRPRQWEA-DSKMSVES 355
Db 71 leekexkvgaeaevalnrmrlleeeleeraeerlkiateklee 115
QY 356 LQNRVTELESVKSGQVARNRTGLLESQLSRHDQMLVHDRLAD 400

RESULT 5
ID R66930 standard; Protein; 885 AA.
AC R66930;
DT 01-SEP-1995 (first entry)
DE AMML chromosome inv(16) product.
KW AMML; acute myelomonocytic leukemia; chromosome-16; inversion;
KW inv(16); CBF-beta; CBF gene; transcription factor; myosin; MYH11;
KW SMHC.
OS Homo sapiens.
FH Key
FH Peptide Location/Qualifiers
FT /label= CBF 1..164
FT Peptide 165..885
FT /label= MYH11
FN W09504067-A.
PD 09-FEB-1995.
PF 26-JUL-1994; U08530.
PR 29-JUL-1993; US-099869.
PA (UNMI) UNIV TEXAS SYSTEM.
PA (TEXA) UNIV TEXAS SYSTEM.
PI Claxton D, Collins FS, Liu P, Siciliano MJ;
DR WPI; 95-082178/11.
DR N-PSDB; 084589.
PT Novel DNA spanning the pericentromeric inversion of chromosome 16 -
PT for the screening of acute myeloid leukaemia
PS Claim 4; Page 34-38; 78pp; English.
CC PCR was performed on total cellular RNA from 5 AMML patients having
CC a pericentromeric inversion of chromosome-16, M450 subtype. Sequencing
CC showed the inv(16) fusion to comprise a sequence from the CBF
CC gene, encoding a novel transcription factor, and the MYH11 gene,
CC encoding smooth muscle myosin heavy chain. In 1 patient, nt 1-492
CC of the CBF gene were fused to nt 994 of MYH11 (shown in

CC Q84589; predicted aa sequence in R66930). Probes based on inv(16)
CC can be used for diagnosis of AMML.
SQ Sequence 885 AA;

Query Match 3.1%; Score 130; DB 12; Length 885;
Best Local Similarity 22.2%; Pred. No. 1.17e-01;
Matches 28; Conservative 38; Mismatches 55; Indels 5; Gaps 5;

Db 204 leaqvqelqskcdgeraraeindkvhlqnevsygmhnaegkaiklakvaslesq 263
QY 269 LEKVSLQNSVSKKNSISLNHQISFELIEIRQKEMLRNNSKIIHLQVIDSQAEK 328
Db 264 lqdtqellqeetrklnvstklq-leeernslqddide-emeaknlerhstlnqls 321
QY 329 LKELDKREIRP-FRQWEADSKMSVSESLQNRVTE-LESVDKSGQ-VARNRTGLLESQLS 385
Db 322 dekkkl 327
QY 386 RHDQML 391

RESULT 6
ID P81187 standard; protein; 1284 AA.
AC P81187;
DT 04-DEC-1990 (first entry)
DE Sequence encoded by a gene fragment coding for major protein of an A-type
DE inclusion body and a promoter region in poxvirus
KW Vaccinia virus; vaccine.
OS Cowpox virus.
FH Key
FH Region Location/Qualifiers
FT /note="Encoded by promoter region"
FN EP-261925-A.
PD 30-MAR-1988.
PF 22-SEP-1987; 308372.
PR 2-SEP-1986; JP-222194.
PR 09-SEP-1987; JP-223972.
PA (TOFU) Toa Nenryo Kogyo KK.
PI Shida H, Funahashi S;
DR WPI; 88-086185/13.
DR N-PSDB; n81538.
PT Gene fragment coding A-type inclusion body in pox virus -
PT used in the construction of recombinant vaccinia viruses for use
PT as vaccines
PS Disclosure; 3-10; 24pp; English.
CC The gene is nonessential for proliferation of poxvirus and is homologous
CC with a corresp. gene of a vaccinia virus and can therefore be used for
CC construction of a recombinant vaccinia virus for use as vaccine. The
CC promoter present upstream of this gene is very strong and is adequate as
CC promoter for expression of an exogenous antigen cell.
SQ Sequence 1284 AA;

Query Match 3.0%; Score 125; DB 1; Length 1284;
Best Local Similarity 22.4%; Pred. No. 2.66e-01;
Matches 22; Conservative 35; Mismatches 36; Indels 5; Gaps 5;

Db 673 lqskdslerlqrcrnateisrlq-yri-tdlerqlndcrnmena-dteremqrlr 729
QY 269 LEKVSLQNSVSKKNSISLNHQISFELIEIRQKEMLRNNSKIIHLQVIDSQAEK 326

Query Match 2.8%; Score 117; DB 11; Length 337;
Best Local Similarity 18.9%; Pred. No. 9.71e-01;
Matches 27; Conservative 47; Mismatches 62; Indels 7; Gaps 7;
Db 130 aeallkkvklfgesngemekdrekldyknkvdadwglireadkireanrlfav 189
QY 266 SNSLEKRVSLQNSVEKNSIQ-SLHNOCSFEIEIEKQEMKLNESKILHLORVDS 324
Db 190 nqkmtalekkkeavegkrqientlkegnldideanrladeinsldyvediqtklpmm 249
QY 325 QAEIKELDKKEIRF-RQNWELADSKMSVSESL-Q-NRVT-BLESDVSAGGVANRITLL 380
Db 250 seelndkiddlsgelkdrkjaek 272
QY 381 ESQLS-RHDDQ-LSVHDIRLADM 401

RESULT 10
ID P90994 standard; protein; 385 AA.
AC P90994;
DT 23-FEB-1990 (first entry)
DE B. subtilis sacuS1 polypeptide
KW levan saccharase
OS Bacillus subtilis.
PN W08909264-A.
PD 05-OCT-1989.
PF 22-MAR-1989; F00134.
PR 22-MAR-1988; FR-003736.
PA (INSP) Inst Pasteur.
PI Kunst F, Debarbouille M, Msadek T, Rapoport G, Klier A, Decorder R,
DR WPI; 89-309330/42.
DR N-PSDB; N91619
PT DNA contg. Bacillus subtilis sacu locus - for inducing overproduction
PT in microorganisms
PS Claim 3; fig. 5; 87pp; French.
CC SacuS1 confers Degg phenotype, and is encoded by part of the 2.55 kb
CC SalI-SphI Sacu locus. This can restore levan saccharase synthesis in
CC B. subtilis sacu- mutants, and overproduces proteins in this or other
CC microorganisms.
SQ Sequence 385 AA;

Query Match 2.8%; Score 117; DB 1; Length 385;
Best Local Similarity 26.5%; Pred. No. 9.71e-01;
Matches 22; Conservative 30; Mismatches 27; Indels 4; Gaps 4;
Db 109 greqlreerddlerllglq-elierselvgqitvlnylnqldrevglldadagkq 167
QY 257 QHVLLEKWSLEKRVSLQNSVEKNSIQSLHNOCSFEIEIEKQEMKLNESKILHLORVDS 315
Db 168 dfgl-rileageeekrvsreth 189
QY 316 -IHLQVRIDSOAEIKELDKKEIR 337

RESULT 11
ID R43563 standard; protein; 476 AA.
AC R43563;
DT 05-APR-1994 (first entry)
DE Hyaluronan receptor.
KW Hyaluronan binding protein; HA; RHAMM; mediated motility; wound;

KW healing; diagnosis; treatment; cell locomotion; tumour invasion;
KW birth defects; inflammatory disorder; Alzheimer's disease; dementia;
KW Parkinson's disease; Huntington's disease; AIDS; diabetes; auto;
KW immune diseases; corneal dystrophy; hypertrophy; surgery; burns;
KW stroke; multiple sclerosis; depression; schizophrenia; CNJ;
KW contraception; in vitro fertilisation; embryo development.
PN W03321312-A.
PD 28-OCT-1993.
PF 13-APR-1993; CA0158.
PR 09-APR-1992; GB-007949.
PA (MANI-) MANITOBA CANCER TREATMENT & RES FOUND.
PA (DTMA-) UNIV.
PA MANITOBA.
PI Turley EA;
DR WPI; 93-351722/44.
DR N-PSDB; 051212.

PT DNA encoding hyaluronan receptor - used to produce proteins and
PT antibodies for alteration of cell locomotion
PS Claim 7; Fig 23; 86pp; English.
CC The sequence is that encoded by a cDNA clone encoding the hyaluronan
CC receptor (HARC). The sequence was obt. by screening a 373 library in
CC lambda gtl with antibodies to HARC. A clone of 1.9 kb was obtained
CC and used to rescreen the library to obtain the full length, 2.9 kb
CC clone. HA is down regulated in stationary normal cells and is only
CC expressed in situations where cell motility is desired, e.g. in
CC wound healing, in response to growth factors and in chemotaxis by
CC white blood cells. HA may be used for diagnosis and treatment of
CC diseases involving cell locomotion, e.g. tumour invasion, birth
CC defects, acute and chronic inflammatory disorders, Alzheimer's and
CC other forms of dementia, AIDS, diabetes, autoimmune diseases, corneal
CC dystrophies and hypertrophies, burns, surgical incisions and adhesions,
CC strokes, multiple sclerosis, depression/schizophrenia related to
CC neuronal growth and pain states involving nerve sprouting; also in CNJ
CC and spinal cord regeneration, contraception, in vitro fertilisation and
CC embryo development.
CC See also R46548-51.
SQ Sequence 476 AA;

Query Match 2.8%; Score 118; DB 8; Length 476;
Best Local Similarity 20.9%; Pred. No. 8.27e-01;
Matches 31; Conservative 41; Mismatches 71; Indels 5; Gaps 5;
Db 101 nllekevelekhargaqaliagelytdags-lrvtaqlesvgekyndagsldv 159
QY 260 NLKEMWSLEKRVSLQNSVEKNSIQSLHNOCSFEIEIEKQEMKLNESKILHLORVDS 318
Db 160 taqlesgekyndagslrvtaqlesgekyndagslrvtaqlesvgekyndagsl 219
QY 319 QRVIDSQAEIKELDKKEIRFQNWELADSKMSVSESLQNRVTELESDVSAGGVANRITLL 377
Db 220 rdvasqlesyksstlkeidkientl 247
QY 378 GILESQLSRA-DQMLS-VHDIRLADM 403

RESULT 12
ID R05766 standard; protein; 462 AA.
AC R05766;
DT 05-NOV-1990 (first entry)
DE Portion of peptide antigen to malarial sporozite.

Query Match	2.7%;	Score 114;	DB 1;	Length 462;
Best Local Similarity	21.9%;	Pred. No. 1.57e+00;		
Matches	32;	Conservative	46;	Mismatches 60; Indels 8; Gaps 8;

```

Dd      262  eqgsedl eqdr lakek iqvgsdler tka steu lgergsd lger irakek lqgsad l e-q 320
Qy      327  EKlKElDKfE-I-RP-FRÖMWEADADSKMSVSSEJLÖNRVTEl SVDRKSAGVARNtGLtESQ 383

Dd      321  errakek lqgsqr d lger lakek i q 346
Qy      384  lSRfDGMlSVHDIrlADMDlGfGVtLE 409

```

ID	RESULT	13
AC	P90954 standard; protein; 769 AA.	
DT	P90954;	
DE	27-Feb-1990 (first entry)	
DE	Yeast topoisomerase I cDNA	
KW	Scleroderma.	
OS	Saccharomyces cerevisiae	
PN	W08909222-A.	
PD	05-OCT-1989.	
PF	22-MAR-1989; 001116.	
PR	23-MAR-1988; 05-172159.	
PA	(RRIG) Brigham and Women's Hospital; (UYJO) John's Hopkins Univ.	
PI	Earnshaw WC, D'Arcy P;	
DR	WPI; 89-309500/42.	
PT	Cloned cDNA encoding eukaryotic topoisomerase I - useful for large scale	
PT	prodn. by recombinant methods	
PS	Disclature; fig. 6; 28pp; English.	
CC	The cDNA of this can be spliced into DNA vectors and used to transform	
CC	hosts for high yield. This has high homology with human topoisomerase I	
SO	Sequence 769 Aa;	
Query Match	2.7%;	Score 111; DB 1; Length 769;

Best Local Similarity 24.7%; Pred. No. 2.55e+00;
Matches 23; Conservative 26; Mismatches 41; Indels 3; Gaps 3;

RESULT	14	
ID	R30729	standard; Protein; 870 AA.
AC	R30729;	
DT	20-MAY-1993	(first entry)
DE	p100 protein from human herpes virus type 6.	
KW	antibody's; monoclonal antibody; ELISA assays; CMV; cytomegalovirus.	
OS	Human herpes virus type 6.	
PN	EP-524421-A.	
PD	27-JAN-1993.	
PF	15-JUN-1992; 110047.	
PR	08-JUL-1991; EP-111338.	
PA	(BEHM) BEHKINGENKE AG.	
PI	Fleckenstein B, Neipel F;	
DR	WPI; 93-028531/04.	
DR	P-RSDb; R30729.	
PT	Human herpes virus type 6 protein p100 DNA sequence - useful in	
PT	prophylaxis, treatment and differential diagnosis of human herpes	
PT	virus-6 infections	
FS	Claim 1; Page 12; 25pp; English.	
CC	This sequence is the p100 protein from human herpes virus type6.	
CC	The protein and antibodies to it can be used for treatment or	
CC	prevention of HHV-6 infections. The DNA , protein and Ab are also	
CC	useful in eg. ELISA assays esp. for differentiating between HHV-6	
CC	and cytomegalovirus infections. These assays are more sensitive	
CC	and specific than immunofluorescence methods currently used.	
SO	Sequence 870 AA;	

Query March	2.7%;	Score 111;	DB 6;	Length 870;
Best Local Similarity	26.4%;	Pred. No. 2.52e+00;		
Matches	32;	Conservative	34;	Mismatches 47;
			Indels	8;
			Gaps	7;
Db	678	grrvnmllsqfntnlisjgkxgledlign-qrgtclklatenksgrseeeanekil-evsn	735	
	:	:	:	:
Qy	319	QRTVID--SQEFKIKEL-DKEIRPFQRMNEEDSKRSSVESLQWRVTELESVDSAGVVAR	375	
	:	:	:	:
Db	736	pqdmfkn-frlqmdlsvspsfrlpdadis-reidsasfkdaidlklpngereidlale	793	
	:	:	:	:
Qy	376	NTGLLESQSLSRHDQMLSVHD-IRLADMDSRGQVLETSYSGVLKIKIRYKRRKQEAVMG	434	
	:	:	:	:
Db	794	k 794		
Qy	435	K 435		
RESULT	15			
ID	R15148	standard; Protein; 475 AA.		
AC	R15148;			
DT	14-FEB-1992	(first entry)		

DE Ro/SSA autoantigen.
KM Autoantibody; autoantigen; SLE; systemic lupus erythematosus.
OS Homo sapiens.
PN M09117171-A.
PD 14-NOV-1991.
PF 07-MAY-1991; U03139.
PR 07-MAY-1990; U5-92070.
PA (OKLA-) OKLAHOMA MED RES FO.
PI Frank MB, Itoh K.
DR WPI: 91-353732/48.
DR N-PSDB; Q14798.
PT DNA encoding an Ro-SSA autoantigen - useful for diagnosing
PT auto-immune disorders or presence of auto-antibodies
PS Disclosure; Fig 2; 41pp; English.
CC A cDNA library (from human thymus mRNA) in lambda gt11 was screened
CC with serum from a patient having systemic lupus erythematosus. Two
CC clones were reactive with sera (from a panel of lupus patients)
CC which contd. autoantibodies against 52 KD protein.
CC Both the cDNA and the protein expressed from it, or portions of it,
CC are useful as diagnostic agents in the identification of patients
CC having autoantibodies and in the identification and analysis of
CC the structural and functional properties of the autoantigen and for
CC application in immunotherapeutic regimens.
SQ Sequence 475 AA;

```

Db      16 cpclidlpfvepvsjlecbhsfgecjsygykgsyvcpcvrcrtfl 59
      | | | | | | | | | | | | | | | | | | | | | | | | |
Oy      53 CEKCHLVLCSPKOTECGHRFCESCAALILSSSSKCTACQESIV 96

Query Match      2.6%; Score 107; DB 3; Length 475;
Best Local Similarity 27.3%; Pred. No. 4.70e+00;
Matches 12; Conservative 13; Mismatches 19; Indels 0; Gaps 0;

Search completed: Tue Mar 19 09:53:45 1996
Job time : 28 secs.

```

Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by IntelliGenetics, Inc.

```

MPsrch_gp  protein - protein database search, using Smith-Waterman algorithm
Run on:      Tue Mar 19 09:52:04 1996;      MacPar time 18.51 Seconds
Tabular output not generated.      775.963 Million cell updates/sec
Title:      >US-08-404-832-2

```

```

Description:      (1-567)  from US0840483z.pdp
Perfect Score:   4162
Sequence:        1 MESKKKMSFGALQJNPFK.....IKDDIIFIKIVIDSJLDPDP 567

Scoring table:   PAM 150
                  Gap 11

Searched:        82306 seqs, 25270970 residues

Post-processing: Minimum Match 0%
                  Listing first 45 summaries

Database:
    pir46
      1:ann1 2:ann2 3:ann3 4:unann1 5:unann2 6:unann3 7:unann4
      8:unann5 9:unann6 10:unann7 11:unann8 12:unrev1
      13:unrev2

Statistics:      Mean 49.231;  Variance 140.106;  scale 0.351

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

```

Result No.	Score	Query length	DB	ID	Description	Pred. No
1	4139	99.4	568 11	A55649	TNR-associated prot	0.00e+00
2	812	19.5	416 11	B55649	TNR-associated prot	2.22e-93
3	191	4.6	458 9	A29361	DG17 protein - alme	9.46e-09
4	164	3.9	760 11	A40195	meprin A (EC 3.4.24.	1.10e-05
5	151	3.6	273 11	A560364	tropomyosin - mgirat	4.63e-04
6	149	3.6	337 11	A47380	RING finger-contain	2.84e-04
7	151	3.6	748 11	S24134	endopeptidase 2 (EC	2.84e-04
8	151	3.6	1969 2	S02771	myosin heavy chain A	2.84e-04
9	145	3.5	284 5	A44860	tropomyosin, oblique	1.23e-03
10	145	3.5	1939 2	A33377	myosin heavy chain,	1.23e-03
11	141	3.4	1961 2	A61231	myosin heavy chain N	3.21e-03
12	142	3.4	1966 2	MMKW	myosin heavy chain B	2.53e-03
13	132	3.2	524972	S24972	tropomyosin alpha, c	2.67e-02
14	132	3.2	507 10	S33068	myosin II heavy chai	2.67e-02
15	132	3.2	760 10	S43883	meprin A (EC 3.4.24.	2.67e-02
16	135	3.2	714 13	S39464	PABA peptide hydrola	1.33e-02
17	135	3.2	1919 2	S03166	myosin heavy chain,	1.33e-02
18	135	3.2	3259 10	S37336	macroglolgin - human	1.33e-02
19	130	3.1	221 5	S05446	tropomyosin, nonmusc	4.24e-02
20	130	3.1	245 13	S24403	tropomyosin alpha -	4.24e-02
21	128	3.1	245 5	C34787	tropomyosin 3 alpha,	6.70e-02
22	128	3.1	248 5	D39816	tropomyosin 5b, fibr	6.70e-02
23	128	3.1	248 5	C39816	tropomyosin 5a, fibr	6.70e-02
24	128	3.1	251 5	B34787	tropomyosin 2 alpha,	6.70e-02
25	131	3.1	257 5	A02991	myosin heavy chain,	3.37e-02
26	128	3.1	284 13	S23470	tropomyosin - Atrica	6.70e-02
27	131	3.1	465 5	A02986	myosin alpha heavy c	3.37e-02
28	129	3.1	944 9	S26710	spindlie pole body pr	5.33e-02
29	127	3.1	961 5	A33626	fibrinogen alpha cha	8.41e-02
30	129	3.1	1053 10	A41642	dynactin - chicken	5.33e-02
31	127	3.1	1078 9	S30722	myosin heavy chain,	8.41e-02
32	128	3.1	1132 9	A53932	hypothetical protein	6.70e-02

33	131	3.1	1509	2	A27224	myosin heavy chain I	3.37e-02
34	129	3.1	1938	2	MMKWL	myosin heavy chain D	5.33e-02
35	130	3.1	1999	2	S21801	myosin heavy chain, 4.24e-02	
36	127	3.1	2007	2	B43402	myosin heavy chain-B	8.41e-02
37	130	3.1	2094	11	S33124	tptr protein - human	4.24e-02
38	129	3.1	2116	5	A26655	myosin heavy chain -	5.33e-02
39	126	3.0	248	5	B30125	tropomyosin beta-3,	1.06e-01
40	126	3.0	621	10	A29965	lamin Dm-0 precursor	1.06e-01
41	126	3.0	676	5	S00084	myosin heavy chain,	1.06e-01
42	125	3.0	1325	13	S16129	dysenin-associated pr	1.32e-01
43	126	3.0	1437	5	S12459	myosin beta heavy ch	1.06e-01
44	126	3.0	1679	9	S48385	hypothetical protein	1.06e-01
45	126	3.0	1937	10	S49478	myosin - human	1.06e-01

ALIGNMENTS

RESULT 1
ENTRY A55649 #type complete
TITLE TNFR-associated protein LMP1 - human
CD40-binding protein
ALTERNATE_NAMES #formal name Homo sapiens #common name man
ORGANISM 23-Mar-1995 #sequence_revision 23-Mar-1995 #text_change
DATE 03-Apr-1995

ACCESSIONS
REFERENCE A55649
#authors Mosialos, G.; Birkenbach, M.; Valamanchili, R.; VandeRade, T.; Mare, C.; Kieff, E.

#journal Cell (1995) 80:389-399
#title The Epstein-Barr virus transforming protein LMP1 engages signaling proteins for the tumor necrosis factor receptor family.

#accession A55649
#molecule_type mRNA
#residues 1-568 #label MOS
#cross-references GB:019260
#note nucleotide sequence not given

REFERENCE A55135
#authors Hu, H.M.; O'Rourke, K.; Boguski, M.S.; Dixit, V.M.
#journal J. Biol. Chem. (1994) 269:30069-30072
#title A novel RING finger protein interacts with the cytoplasmic domain of CD40.

#accession A55135
#molecule_type mRNA
#residues 1-133,135-404, 'G', 406-568 #label HDA
#cross-references GB:015637
#note nucleotide sequence not given
#note coiled coil

KEYWORDS #region RING-finger motif
FEATURE #molecular-weight 64490 #checksum 8660
SUMMARY #length 568

Query Match 99.4%; Score 4139; DB 11; Length 568;
Best Local Similarity 99.6%; Pred. No. 0.00e+00;
Matches 566; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
Db 1 mesekndspalgtlppkllhdsrsgtvpfpeqgykktvktvckekchyl 60
Qy 1 MESKKNDSPALGTLPPLKLLHDSRSGTVPFPEQGYKKTVKTVEKCKECHYL 60

Db	61	csptqteqhnrfcesmaallssespkctacgesiykdvfkdnockreialqiyene	120
Qy	61	CSPTQTEQHNFCESCMAALLSSSPKCTACGESIYKDVFKDNCKREIALQIYENE	120
Db	121	argaaeqmqlghllvhlkndcheelpcvppdkkvlkldrhvckackyreacahc	180
Qy	121	ARGAAEQMLGHL-VHLKNDCHLELPCVPDPKKVLKLDLHDVCKACKYREATCSHC	179
Db	181	ksqymalqkhdtdcpcvvascpkcsvqtlireelsahlsecvnapstcsfrkygc	240
Qy	180	KSQYPMALQKHEDTDCPCVVASCPKCSVQTLIRELSAHLSECVNAPSTCSFRKYCV	239
Db	241	fgtngqtkahaaasavqvnllkewanslekksvllqnesveksksjqlmqicse	300
Qy	240	FGTNQOIKAHAASSAVQVNNLLKEWNSLEKKSILLQNESEKKSISQSLNQICSEI	299
Db	301	eierqemlrneekihlgrvidsqaeikeldkelpfrqmwesadmkssveslgr	360
Qy	300	EIERQEMLRNESHILHLPVIDSQAEIKELDKELRFRQMWESADMSKSSVESLGR	359
Db	361	vtelssvdksagvarntglleeqlerhdqmlsvhdirladmlrfqvletasyngvliw	420
Qy	360	VTELESVDKSAQVARNLTGLLESQLSRHDQMLSVHDIRLADMLRFGVLETASVNGVLIW	419
Db	421	kirdykrkqeaumgtclsiysqpfyfygkmcarylmgdqmgkgtahslffvintg	480
Qy	420	KIRDYKRKQEAUMGKTLISYSPFYGYGKMCARYLMDGMDGKGTHTSLFFVINTG	479
Db	481	eydallpwpfkykvrlmldmgssrhlhgafkfpdpnasfkkprgmmlaagcvfvaq	540
Qy	480	EYDALLPWPFKQVRLMLDMDGSSRHLHGAFKFPDPNSSFKKPRTEAMIASGCVFVAQ	539
Db	541	tvlenqykddtffikvfvtdslpdp	568
Qy	540	TVLENGYIKDDTFFIKVIVDTSIDPDP	567

RESULT 2
ENTRY B55649 #type complete
TITLE TNFR-associated protein EB16 - human
ORGANISM #formal name Homo sapiens #common name man
DATE 23-Mar-1995 #sequence_revision 23-Mar-1995 #text_change

ACCESSIONS B55649
REFERENCE B55649
#authors Mosialos, G.; Birkenbach, M.; Valamanchili, R.; VandeRade, T.; Mare, C.; Kieff, E.

#journal Cell (1995) 80:389-399
#title The Epstein-Barr virus transforming protein LMP1 engages signaling proteins for the tumor necrosis factor receptor family.

#accession B55649
#status preliminary
#molecule_type mRNA
#residues 1-416 #label MOS
#cross-references GB:019261

SUMMARY #length 416 #molecular-weight 46163 #checksum 6815
Query Match 19.5%; Score 812; DB 11; Length 416;
Best Local Similarity 38.9%; Pred. No. 2.29e-93;

ENTRY	5	6
TITLE	A60364	#type complete
ORGANISM	tropomyosin - migratory locust	
DATE	#formal_name Locusta migratoria #common_name migratory locust 03-Feb-1993 #sequence_revision 03-Feb-1993 #text_change 18-Jun-1993	
ACCESSIONS	A60364	
REFERENCE	A60364	
#author	Kriegler, J.; Ramling, K.; Knipfer, M.; Grau, M.; Mertens, S.; Beer, H.	
#journal	Insect Biochem. (1990) 20:1173-1184	
#title	Cloning, sequencing and expression of locust tropomyosin. A60364	
#accession	not compared with conceptual translation	
#status		
#molecule_type	mRNA	
##residues	1-283	##label KRI
CLASSIFICATION	#superfamily tropomyosin	
KEYWORDS	coiled coil; hepled repeat	
SUMMARY	#length 283; #molecular-weight 32439 #checksum 4917	
Query Match	3.6%; Score 149; DB 5; Length 283;	
Best Local Similarity	19.0%; Pred. No. 4,63e-04;	
Matches	20; Conservative	36; Mismatches 46; Indels 3; Gaps 3;
Db	13 lekdnaldrallceqgdanl-ra-ekaeearalqkrlqtendldqtgeslqymak 70	
Oy	297 FEIEIRKQEMLRNNESEKILRLORVIDSQAEKLEKDEKIRPFQWEEA-DSMKSYES 355	
Db	71 leekelagaaeseaalnrirldleedlargeelataataakae 115	
Oy	356 LQNRVTELESYDKSAGVARNRTGILLESQLSRHDQMLSVHDRIAD 400	
ENTRY	6	
TITLE	A47380	#type complete
ORGANISM	RING finger-containing DNA binding protein RING1 - human	
DATE	#formal_name Homo sapiens #common_name man 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 18-Nov-1994	
ACCESSIONS	A47380	
REFERENCE	A47380	
#author	Lovring, R.; Hanson, I.M.; Borden, K.L.; Martin, S.; O'Reilly, N.J.; Evan, G.I.; Rahman, D.; Pappin, D.J.; Trowadeale, J.; Freemont, F.S.	
#journal	Proc. Natl. Acad. Sci. U.S.A. (1993) 90:2112-2116	
#title	Identification and preliminary characterization of a protein motif related to the zinc finger.	
#cross-references	MUID:93211912	
#accession	A47380	
##status	preliminary	
##molecule_type	DNA; protein	
##residues	1-377	##label LOV
##cross-references	NCBIN:128010; NCBIPI:128011	
#note	sequence extracted from NCBI backbone	
SUMMARY	#length 377 #molecular-weight 39145 #checksum 7895	
Query Match	3.6%; Score 151; DB 11; Length 377;	
Best Local Similarity	24.7%; Pred. No. 2.84e-04;	
Matches	24; Conservative	30; Mismatches 39; Indels 4; Gaps 4;

Db	11	relhslmelipidmlkntmtcklrfcdscivaltalrsgmkpcptckklivktrslrp	70
Oy	45	KIVEDKYKKEKHVLVCSFKOT-ECGHRFCSCMAALISSSPKCTACQESTIVKDKVTKD	103
Db	71	dp-nfda1sklypsreeyeahqdrvl1-flsr1hng	105
Oy	104	NCKRETLALQIY-CRNESRCGACQIMLGHVHKND	139
RESULT	7		
ENTRY			
TITLE	S24134	#type complete	
ALTERNATE_NAMES	endopeptidase 2 (EC 3.4.24.-) - rat		
ORGANISM	endopeptidase 24.18	#formal name Rattus norvegicus	#common name Norway rat
DATE	02-Dec-1993	#sequence_revision 01-Sep-1995	#text_change 01-Sep-1995
ACCESSIONS	S24134		
REFERENCE	S24134		
#authors	Corbelle, D.; Gaudoux, F.; Mainwright, S.; Ingram, J.; Kenny A. J.; Bollean, G.; Crine, P.		
#journal	FEBS Lett. (1992) 309:203-208		
#title	Molecular cloning of the alpha-subunit of rat endopeptidase-24.18 (endopeptidase-2) and co-localization with endopeptidase-24.11 in rat kidney by in situ hybridization.		
#cross-references	MFID:92371675		
#accession	S24134		
#status	preliminary		
#molecule	type mRNA		
#residues	1-748	#label COR	
FEATURE	hydroxylase; metalloprotein; proteinase; zinc		
156,160,166			
SUMMARY			
	#binding site zinc (His) #status predicted		
	#length 748 #molecular-weight 85138 #checksum 2333		
Query Match	3.6%; Score 151; DB 11; Length 748;		
Best local similarity	28.9%; Pred. No. 2.84e-04;		
Matches	26; Conservative 23; Mismatches 37; Indels 4; Gaps 4;		
Db	436	wctmrnsgvltvktgdt1-v-sprfyne-ygfgfvtlpyngritsngy1glaflny	492
Oy	418	IKRIIRYKRRKEAVMGKTLISYSPFYTGFGKMCARVYINGDMGKTHLSFVIM	477
Db	493	sgndv1llepvenegaimtlldgdpdrn	522
Oy	478	RGEYDALIPFKQKVTIM-IMDGSSSRH	506
RESULT	8		
ENTRY	S02771	#type complete	
TITLE	myosin heavy chain A - Caenorhabditis elegans		
CONTAINS	myosin ATPase (EC 3.6.1.32)		
ORGANISM	#formal name Caenorhabditis elegans		
DATE	31-Dec-1993	#sequence_revision 31-Dec-1993	#text_change 27-Jan-1995
ACCESSIONS	S02771		
REFERENCE	S02771		
#authors	Dibb, N. J.; Maruyama, I. N.; Krause, M.; Karn, J.		
#journal	J. Mol. Biol. (1989) 205:603-613		
#title	Sequence analysis of the complete Caenorhabditis elegans		

		myosin heavy chain gene family.	
#cross-references		MIMD:89178677	
#accession	S02771		
#molecule	_type DNA		
#residues	1-1969 ##label D1B		
#cross-references	EMBL:X08067		
#note	the nucleotide sequence is not given		
GENETICS		myo-3	
#gene	4671; 116/3; 169/1; 269/1; 445/2; 1896/3		
#introns			
CLASSIFICATION		#superfamily myosin heavy chain; myosin head homology	
KEYWORDS		actin binding; ATP; coiled coil; hydrolase; methylation;	
		muscle contraction; tandem repeat	
FEATURE			
89-779	#domain myosin head homology #label HEA\		
179-186	. #region nucleotide-binding motif A (P-loop) \		
667-689	#region actin-binding #status predicted\		
770-784	#region actin-binding #status predicted\		
852-1969	#domain coiled coil #status predicted #label COI\		
852-1166	#region S2\		
1167-1969	#region light meromyosin\		
130	#modified site N6,N6,N6-trimethyllysine (lys) #status predicted\		
185	#binding site ATP (lys) #status predicted\		
707, 717	#active site Cys #status predicted		
SUMMARY	#length 1969_	#molecular_weight 225508	#checksum 6892
Query Match 3.6%; Score 151; DB 2; Length 1969;			
Best Local Similarity 21.2%; Pred. No. 2,84e-04;			
Matches 28; Conservative 49; Mismatches 47; Indels 8; Gaps 7;			
Db	985	dnnrlrsiqdmanqdeavklnkek-khgeenrklnejdgsedkvnhlekinklegq	1043
Oy	270	EKKVSLIÖNSVEKNKSIÖSLIHÖISFELEIFRÖ-KEMLRNNEKSIILHÖRVIDSÖAK	328
Db	1044	mdeleenidreksrgdiekarkvveg-dlkvaq-endeltkg-kh-dv-etlkrke	1097
Oy	329	LKEIDKEIRPFROWEADSMKSSVESLÖNRVTELESVDKSGAVARMTGLLESQLSRHD	388
Db	1098	edlnhtnakiiae 1109	
Oy	389	ÖMISVHDIRLAD 400	
ENTRY	9	A44980	#type complete
TITLE	tropomyosin, obliquely striated muscle - nematode		
ORGANISM	(Trichostrongylus colubriformis)		
DATE	#formal_name Trichostrongylus colubriformis		
ORGANISM	28-Apr-1993 #sequence_revision 28-Apr-1993 #text_change 12-Apr-1995		
ACCESSIONS	A44980		
REFERENCE	A44980		
#authors	Frenkel, M.J.; Savin, K.W.; Bakker, R.E.; Ward, C.W.		
#journal	Mol. Biochem. Parasitol. (1989) 37:191-199		
#title	Characterization of cDNA clones coding for muscle tropomyosin of the nematode Trichostrongylus colubriformis.		
#accession	A44980		
#status	preliminary		
#molecule_type	mRNA		

	Query Match	3.5%; Score 145; DB 5; Length 284;	
	Beat Local Similarity	21.3%; Prec. No. 1,23e-03;	
	Matches	26; Conservative	49; Mismatches 41; Indels 6; Gaps 6;
Db	85	vaalnrrmtllleeleraerikatekleatnhd-eeseryrk-menisfgde-era	141
Oy	287	IQSIHNOICFEIETREKEMRNNSKIILHOVIDSOAEKLEKEIDRFRONWEA	346
Db	142	nrlaelqkeag-mlae-ea-dtkydevarklamreadleraeeraaagnkivleeeir	198
Oy	347	DSKSSVESLÖNRVTELESVDKSACQVARNITGLESQSLSRDMLSVHDRIADLDGFQ	406
Db	199	vv 200	
Oy	407	VL 408	
RESULT	10		
ENTRY	A33977	#type complete	
TITLE	myosin heavy chain, cellular - chicken		
CONTAINS	myosin ATPase (EC 3.6.1.32)		
ORGANISM	#formal name Gallus gallus #common name chicken		
DATE	31-Dec-1993 #sequence_revision 31-Dec-1993 #ext_change 27-Jan-1995		
ACCESSIONS	A33977		
REFERENCE	A33977		
#authors	Shoher, R.V.; Conti, M.A.; Kawamoto, S.; Preston, Y.A.; Brill, D.A.; Adelstein, R.S.		
#journal	Proc. Natl. Acad. Sci. U.S.A. (1989) 86:7726-7730		
#title	Cloning of the cDNA encoding the myosin heavy chain of a vertebrate cellular myosin.		
#cross-references	MOTID:90046668		
#accession	A33977		
#molecule_type	mRNA		
#residues	1-1959	#label SHO	
#cross-references	GB:M26510		
CLASSIFICATION	#superfamily myosin heavy chain; myosin head homology actin binding; ATP; coiled coil; hydrolase; methylation; tandem repeat		
KEYWORDS			
FEATURE			
84-764	#domain myosin head homology #label HEA\		
174-181	#region nucleotide-binding motif A (P-loop)\		
352-365	#region actin-binding #status predicted\		
626-640	#region actin-binding #status predicted\		
837-1936	#domain coiled coil #status predicted #label COI\		
837-1277	#region S2\		
1278-1959	#region light meromyosin\		
1937-1959	#domain carboxyl-terminal #label CBT\		
125	#modified site N6,N6,N6-trimethyllysine (Lys) #status predicted\		
180	#binding site ATP (Lys) #status predicted\		
694, 704	#active site Cys #status predicted		
SUMMARY	#length 1959 #molecular-weight 226502 #checksum 3641		
Query Match	3.5%; Score 145; DB 2; Length 1959;		

[illegible]

180	694, 704	SUMMARY	predicted\ #binding site ATP (Lye) #status predicted\ #active_site Cys #status predicted #length 1961 #molecular-weight 226741 #checksum 9108
Query Match	3.4%;	Score 141;	DB 2; Length 1961;
Best Local Similarity	27.8%;	Pred. No. 3,21e-03;	
Matches	32;	Conservative 26;	Mismatches 48; Indels 9; Gaps 7;
Db	1461	yaeeedraaaareketklsrlarlsaeameqkaelerlnkqfr-t--emedlmskadv	1517
Oy	297	FEIEIERKEMENNNESKILHQRVIDSDQAKKEIKDKERIPRPNWEADSNKSSVBSL	356
Db	1518	gksvhelektraleqveamktyleledatgedakrlr-ewlqgsmkafqe	1571
Oy	357	QNRVTELESVDKS-AGQVA-RNTGL--LESOL-SRHDQMLSVHDIPLADMDLGFQ	406
RESULT	12		
ENTRY	MMKM	#type complete	
TITLE	myosin heavy chain B - Caenorhabditis elegans		
CONTRINS	myosin ATPase (EC 3.6.1.32)		
ORGANISM	#journal name Caenorhabditis elegans		
DATE	13-Jun-1983 #sequence_revision 19-Feb-1984 #text_change 27-Jan-1995		
ACCESSIONS	A93958; A93287; A02992		
REFERENCE	A93958		
#authors	Karn, J.; Brenner, S.; Barnett, L.		
#journal	Proc. Natl. Acad. Sci. U.S.A. (1983) 80:4253-4257		
#title	Protein structural domains in the Caenorhabditis elegans unc-54 myosin heavy chain gene are not separated by introns.		
#cross-references	M0ID:83273600		
#accession	A93958		
#molecule_type	DNA		
#residues	1-1966	#label	KAR
REFERENCE	A93287		
#authors	McLachlan, A.D.; Karn, J.		
#journal	Nature (1982) 299:226-231		
#title	Periodic charge distributions in the myosin rod amino acid sequence match cross-bridge spacings in muscle.		
#cross-references	M0ID:82272395		
#accession	A93287		
#molecule_type	DNA		
#residues	850-1336, 'R', 1338-1879, 'L', 1881-1966	#label	MCL
GENETICS			
#gene	unc-54		
#introns	21/3; 64/2; 114/3; 267/1; 528/3; 1750/3; 1822/3; 1897/3		
CLASSIFICATION	#superfamily myosin heavy chain; myosin head homology		
KEYWORDS	actin binding; ATP; coiled coil; hydrolase; methylation; muscle contraction; tandem repeat		
FEATURE			
87-778	#domain myosin head homology	#label	HEA
177-184	#region nucleotide-binding motif A (P-1loop)		
665-687	#region actin-binding	#status	predicted
769-783	#region actin-binding	#status	predicted
851-1966	#domain coiled coil	#status	predicted
831-1165	#region 52\		
1166-1966	#region light meromyosin		

128		#modified site N6,N6,N6-trimethyllysine (Lys) #status predicted\
183		#binding site ATP (Lys) #status predicted\
705,715		#active site Cys #status predicted
SUMMARY		#length 1966 #molecular-weight 225125 #checksum 6382
Query Match	3.4%	Score 142; DB 2; Length 1966;
Best Local Similarity 24.2%;		Pred. No. 2.53e-03;
Matches 32;		Conservative 37; Mismatches 55; Indels 8; Gaps 8;
Db	984	dhqjrlsldgmvgqddaaiklnkex-khqeelnrklnmedlqseedkgnbnkvakleqt 1042
Oy	270	EKKVSLIÖNÖSVEKNKSIGSLIÖNÖCSPEIEIERÖ-KEMIRNNESKILHÖGVDSÖAEK 328
Db	1043	lddledlerekarardldkqrkveg-elkiag-enides-gr-qrh-dl-ennlkke 1036
Oy	329	IKELDKKIRPRÖNWEADSMKSSVSLÖNÖVTELESVDKSGAGVARNÖGLESÖLSRD 388
Db	1097	selhvasarld 1108
Oy	389	ÖMISVHDIRLAD 400
RESULT 13		
ENTRY	S24972	#type complete
TITLE	tropomyosin alpha, cardiac - pig	
ORGANISM	#formal name Sus scrofa domestica #common name domestic pig	
DATE	20-Feb-1995	#sequence_revision 20-Feb-1995 #text_change 12-Apr-1995
ACCESSIONS	S24972	
REFERENCE	S24972	
#author	Whitby, F.G.; Kent, H.M.; Stewart, F.; Stewart, M.; Xie, X.; Hatch, V.; Cohen, C.; Phillips Jr., G.	
#submission	submitted to the EMBL Data Library, April 1992	
#description	Structure of tropomyosin at 9 Angstroms resolution.	
#accession	S24972	
##status	preliminary	
##molecule_type	mRNA	
##residues	1-284	##label WHI
##cross-references	EMBL:X66274	
CLASSIFICATION	#superfamily tropomyosin	
SUMMARY	#length 284 #molecular-weight 32729	#checksum 9163
Query Match	3.2%	Score 132; DB 5; Length 284;
Best Local Similarity 23.9%;		Pred. No. 2.67e-02;
Matches 27;		Conservative 39; Mismatches 39; Indels 8; Gaps 6;
Db	45	elqkklkatedeldkysealkdaqeklelaekkatdaead-vaslnrriqif-e--eeld 100
Oy	289	SLHÖNÖCSPEIEIERÖKEMIRNNESKI-LIHÖPVDSÖAEKIKELDKKIRPRÖNWEAD 347
Db	101	--taq-erlatalqkleekaaadsergmkviasirzqkdeekmeiqelqke 150
Oy	348	SMKSSVSLÖNÖVTELESVDKSGAGVARNÖGLESÖLSRDÖMISVHDIRLAD 400
RESULT 14		
ENTRY	S33068	#type complete
TITLE	myosin II heavy chain - fluke (Schistosoma mansoni)	
ORGANISM	#formal name Schistosoma mansoni	

DATE	22-Nov-1993 #sequence_revision 26-May-1995 #text_change 26-May-1995					
ACCESSIONS	S33068					
REFERENCE	A46514					
#authors	Solsson, L.M.A.; Masterson, C.P.; Tom, T.D.; McNally, M.T.;					
#journal	J. Immunol. (1992) 149:3612-3620					
#title	Induction of protective immunity in mice using a 62-kDa recombinant fragment of a Schistosoma mansoni surface antigen.					
#cross-references	MOTID:93056536					
#accession	S33068					
##status	preliminary					
##molecule_type	mRNA					
##residues	1-527 ##label SOI					
##cross-references	EMBL:X65351					
##note	the authors translated the codon CAA for residue 346 as lys					
SUMMARY	#length 527	#molecular_weight 61535	#checksum 5188			
Query Match	3.2%;	Score 132;	DB 10;	Length 527;		
Best Local Similarity	24.4%;	Pred. No. 2.67e-02;				
Matches	29;	Conservative	33;	Mismatches	51;	
				Indels	6;	
				Gaps	4;	
D6	255 dneurtlgseamagdegmklkkdkklnleekrtglaqaeakvnhlkkaleet1 314 ::: : EKVSYLLQNSVEYKKNKSLOSLSHQCSEFIEIEKGKEMLNNNSSKIHLQRIDSQAKL 329					
D6	315 demeenlareqtkirgvekskrkleq-dlkatq-etvd-dlervkrd--leeqlrke 367 : :: KEIDKEIRPROMWEADSKMSVSYSIONRVTELESVDKSGAGVANRTGLLESQSLRAD 388					
RESULT	15					
ENTRY	S49383	#type complete				
TITLE	meprin A (EC 3.4.24.18) - human					
ORGANISM	16-Feb-1995 Homo sapiens [common name man					
DATE	26-May-1995 #sequence_revision 12-May-1995 #text_change					
ACCESSIONS	S49383					
REFERENCE	S49383					
#authors	Eldering, J.A.; Groenberg, J.; Sterchi, E.E.					
#submission	submitted to the EMBL Data Library, September 1994					
#description	Cloning and the PABA-peptide hydrolase beta subunit: coexpression is required for plasma membrane localization of the alpha subunit in COS-1 cells.					
#accession	S49383					
##status	preliminary					
##molecule_type	mRNA					
##residues	1-700 ##label EMD					
##cross-references	EMBL:X61333					
CLASSIFICATION	#superfamily MAM homology; astacin homology					
FEATURE						
11-257	#domain astacin homology #label AST\					
260-429	#domain MAM homology #label MAM					
SUMMARY	#length 700	#molecular_weight 79458	#checksum 916			
Query Match	3.2%;	Score 132;	DB 10;	Length 700;		
Best Local Similarity	33.3%;	Pred. No. 2.67e-02;				
Matches	30;	Conservative	20;	Mismatches	29;	
				Indels	11;	
				Gaps	7;	

OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 CC EUTHERIA; RODENTIA.
 [1]
 RP SEQUENCE FROM N.A.
 RM 94349371
 RA ROTHE M., WONG S.C., HENZEL W.J., GOEDEL D.V.;
 RL CELL 78:681-692(1994).
 CC -1- FUNCTION: SIGNAL TRANSDUCER ASSOCIATED WITH THE CYTOPLASMIC DOMAIN
 OF THE 75 KD TUMOR NECROSIS FACTOR RECEPTOR (TNF-R2).
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
 CC -1- SUBUNIT: HETERODIMER OF TRAF1 AND TRAF2.
 CC -1- SIMILARITY: CONTAINS A C3HC4-CLASS ZINC FINGER.
 DR EMBL; L33303; MTRAF2A.
 KW ZINC-FINGER; COILED COIL.
 FT ZN_FING 34 72 C3HC4-TYPE.
 SQ SEQUENCE 501 AA; 56026 MM; 1264825 CN;
 Query Match 19.2%; Score 798; DB 7; Length 501;
 Best Local Similarity 52.7%; Pred. No. 5.03e-123;
 Matches 98; Conservative 49; Mismatches 37; Indels 2; Gaps 2;
 Db 317 lantvqqlersjgkldiamadlegkvelevstvgvfiwkadfrkxgeavgrtai 376
 QY 380 LESQLSRHDQMLSVHDIRLADMDIGFOVLETAASNGVLIWKIDRRKQEAVMGKTLST 439
 Db 377 fspafyerygkncilvylngdgtgrghlsifvwmkqpdallqwpfnqgvtlmld 436
 QY 440 YSOFYTGFGYKRCARVYLNCGMGKTHLSFVIMRGEVDALLPWFKQVTLMLMD 499
 Db 437 h-nurehvidafprcdsssfqrpvsdmiasecpifcpvskmeaknsyvdafikai 495
 QY 500 QGSSRRHLGDAFKEDPNSSSFKKPTGEMNIAAGCPVFAQTGLE-NGTVIKDDTIFIKVI 558
 Db 496 vdlgtl 501
 QY 559 VDTSDL 564
 RESULT 2
 ID TRF1 MOUSE STANDARD; PRT; 409 AA.
 AC P39428;
 DT 01-FEB-1995 (REL. 31, CREATED)
 DT 01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)
 DT 01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)
 DE TNF RECEPTOR ASSOCIATED FACTOR 1 (TRAF1).
 OS MUS MUSCULUS (MUSE).
 CC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 CC EUTHERIA; RODENTIA.
 [1]
 RN SEQUENCE FROM N.A., AND SEQUENCE OF 123-135 AND 390-402.
 RM 94349371
 RA ROTHE M., WONG S.C., HENZEL W.J., GOEDEL D.V.;
 RL CELL 78:681-692(1994)
 CC -1- FUNCTION: SIGNAL TRANSDUCER ASSOCIATED WITH THE CYTOPLASMIC DOMAIN
 OF THE 75 KD TUMOR NECROSIS FACTOR RECEPTOR (TNF-R2).
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
 CC -1- SUBUNIT: HETERODIMER OF TRAF1 AND TRAF2.
 DR EMBL; L33302; MTRAF1A.
 KW COILED COIL.
 SQ SEQUENCE 409 AA; 45464 MM; 857796 CN;

Query Match 19.1%; Score 796; DB 7; Length 409;
 Best Local Similarity 38.8%; Pred. No. 1.22e-122;
 Matches 149; Conservative 94; Mismatches 117; Indels 24; Gaps 18;
 Db 40 lrdedctricpctadn-lhvspgsplgtg-kh-sdvaeeaimcpfagvgsfsgspqs 96
 QY 188 LQKHEDDTCF-CVAVSCPHHCVSQVTLRLSELASHLSECVNAPESTCFKRYGCVFGOTNOQ 246
 Db 97 mgeheatsqeshlylllavlkewkspsgsnlgsapm-alernulselqgaaveatgdlv 155
 QY 247 IKAHDAASSAVOHVLL-----KENSNSLEKRVSLQNESVEKRSIOSLNOI-CSEFEI 301
 Db 156 dcyrapceesqee-lalqlvk-e-kllaqlkeklvfv-an-lva-vlnkevea--shla 207
 QY 302 ERQKEMLNNSKSLIHLQVVIDSQAEKIKEDKEIRPFRQNMEDASMSKSVESLQNNVT 361
 Db 208 laaslngs-qldrehlslslegvvelqgtlaqkdqvgiklshslimeeasfdgtflwk 265
 QY 362 ELESVDKSAAGVARNHCTGL-LESQLSRHDQMLSVHDIRLADMDIGFOVLETAASNGVLIWK 420
 Db 266 ltnvtrkheevcgrtvalsfpafyrtakvgkicliylngdsgokthlsfviymrge 325
 QY 421 IROTKRRKQEAVMGKTLSTLSQPTTYFGTRKCARVYLNCGMGKTHLSFVIMRGE 480
 Db 326 ydallpwpfrnkvtfmldq-nurehaidafprldesaeqrpqseetnaasecpifrlps 384
 QY 481 YDALLPWFKQKQVTLMLMDQGSSRRHLGDAFKEDPNSSSFKKPTGEMNIAAGCPVFAQT 540
 Db 385 klgsphaykvkdtmfikciydr 408
 QY 541 VLENG--TYIKDDTIFIKVIDTS 562
 RESULT 3
 ID DG17 DICD1 STANDARD; PRT; 458 AA.
 AC P11467;
 DT 01-OCT-1989 (REL. 12, CREATED)
 DT 01-OCT-1989 (REL. 12, LAST SEQUENCE UPDATE)
 DT 01-FEB-1994 (REL. 28, LAST ANNOTATION UPDATE)
 DE DG17 PROTEIN.
 GN ZFAA OR DG17.
 OS DICTYOSTELIUM DISCOIDEUM (SLIME MOLD).
 CC EUKARYOTA; PROTOZOA; SARCOMASTIGOPHORA; RHIZOPODA;
 CC EUMYCETIZOA; DICTYOSTELIA.
 [1]
 RN SEQUENCE FROM N.A.
 RM 88142840
 RA DRISCOLL D.M., WILLIAMS J.G.;
 RL MOL. CELL. BIOL. 7:4482-4489(1987).
 CC -1- THE EXPRESSION OF DG17 PROTEIN IS DEVELOPMENTALLY REGULATED.
 CC -1- INDUCTION: BY CAMP DURING AGGREGATION.
 DR PIR; A29361; A29361.
 DR DICTYB; DD02010; ZFAA.
 KW DEVELOPMENTAL PROTEIN; ZINC-FINGER.
 FT ZN_FING 25 67 TO DROSOPHILA SINA (AA 71-108).
 FT ZN_FING 27 66 POTENTIAL.
 FT ZN_FING 178 198 POTENTIAL.
 SQ SEQUENCE 458 AA; 53015 MM; 1107348 CN;

```

Overy Match Similarity 4.6%; Score 191; DB 2; Length 458;
Best Local Similarity 25.8%; Pred. No. 1.82e-12;
Matches 39; Conservative 36; Mismatches 63; Indels 11; Gaps 10.

Db 124 kdeengcckekievcdgidshlhn-cqykfvrcsfkyc-ekllr--m-nslkmngfklytc 178
::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Qy 118 RNSRGCAGEDLMGHL-VHLKNDCHFEELPCRPDCKEYKLRDLBDHYEKACKYREATC 176

Db 179 dfctrdidkketletytkr-cmpripdcsqcsvk-ietkslihdncctctgpclyf 236
::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Qy 177 SHCK-SQVPMIALQKHEIDTCGCVVAVSCPHEKCSVQTLLRELSAHL-SECVNAPSTCSFK 234

Db 237 eggckvcmkrseqlnhlerlvhngymglile 267
||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Qy 235 RYCGVCGTNGQIKAEHASSAVQH-VNLKE 264

RESULT 4
ID MEPA MOUSE STANDARD; PRT; 760 AA.
AC P28825;
DT 01-DEC-1992 (REL. 24, CREATED)
DT 01-DEC-1992 (REL. 24, LAST SEQUENCE UPDATE)
DT 01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)
DE MEPRIN A ALPHA-SUBUNIT PRECURSOR (EC 3.4.24.18) (ENDOPEPTIDASE-2)
DS (MEP-1).
OS MOS MUSCULUS (MOUSE).
OC EKKAROTIA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; RODENTIA.
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=C57BL/6, AND C3H/HE; TISSUE=KIDNEY;
RM 92250517
RA JIANG W., GORBEA C.M., FLANNERY A.V., BEYNON R.J., GRANT G.A.,
RA BOND J.S.;
RN J. BIOL. CHEM. 267:9185-9193(1992).
RN [2]
RP SEQUENCE OF 77-275 FROM N.A.
RM 92042028
RA DOMENROTCH E., STERCHI E.E., JIANG W., WOLZ R.L., BOND J.S.,
RA FLANNERY A.V., BEYNON R.J.;
RN J. BIOL. CHEM. 266:21381-21385(1991).
RN [3]
RP CHARACTERIZATION.
RM 91355206
RA WOLZ R.L., HARRIS R.B., BOND J.S.;
RL BIOCHEMISTRY 30:8488-8493(1991).
CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF PROTEIN AND PEPTIDE SUBSTRATES
CC PREFERENTIALLY ON CARBOXYL SIDE OF HYDROPHOBIC RESIDUES.
CC -1- COFACTOR: ZINC.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- SUBUNIT: HOMOTETRAMER OF ALPHA OR BETA SUBUNITS; HETEROTETRAMER
CC OF TWO ALPHA AND TWO BETA SUBUNTS ARE FORMED BY NON-COVALENT
CC ASSOCIATION OF TWO DISULFIDE-LINKED HETERODIMERS; GENETIC FACTORS
CC DETERMINE WHICH OLIGOMER(S) WILL BE FORMED (STRAIN-SPECIFIC).
CC -1- PTM: N-G-LINKOSYLATED; AT LEAST 3 OF THE POTENTIAL SITES ARE USED.
CC -1- TISSUE SPECIFICITY: KIDNEY, INTESTINAL BRUSH BORDERS, AND
CC SALIVARY DUCTS.
CC -1- SIMILARITY: HIGH TO OTHER ZINC METALLOPROTEASES. BELONGS TO THE
CC ASTHININ SUBFAMILY.

```

[illegible]

Query Match	3.6%;	Score 149;	DB 7;	Length 283;
Best Local Similarity	19.0%;	Pred. No. 2.69e-06;		
Matches	20;	Conservative	36;	Mismatches 46;
			Indels	3;
			Gaps	3;

Db 13 lehmaldallieggardann-rx-ekaaeeaaalkkrtigtlendldgtsgslqymak 70
 297 FELEIERKÖMELRNNESSKRIHLHÖRVIDSÖAKKIKEIDKELRPRÖUMWELA-DSMKS SVES 355
 71 leekakaignaeevaalnrirqlleedlerseeatlatataklae 115
 356 lönrrtellesvkrsgoavarnngtelleslshdmoelshvdhrriad 400

RESULT	6	STANDARD;	PRT;	377 AA.
ID	RINI HUMAN			
AC	006587;			
DT	01-JUN-1994 (REL. 29, CREATED)			
DT	01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)			
DT	01-OCT-1994 (REL. 30, LAST ANNOTATION UPDATE)			
DE	RING1 PROTEIN.			
GN	RING1.			
OS	HOMO SAPIENS (HUMAN).			
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;			
OC	EUHETERIA; PRIMATES.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RM	93211912			
RA	LOVEING R., HANSON I. M., BORDEN K. L. B., MARTIN S., O'REILLY N. J.,			
RA	EVAN G. I., RAHAM D. I., PAPEIN D. J. C., THOMASDALE J., FREEMONT P. S.;			
RL	PROC. NATL. ACADE. SCI. U.S.A. 90:2112-2116(1993).			
CC	-1- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).			
CC	-1- SIMILARITY: CONTAINS A C3HC4-CLASS ZINC FINGER.			
EMBL	214000; HSRING1.			
DR	PIR; A47380; A47380.			
DR	HSP; P28990; 1CHC.			
DR	PROSITE; P500518; ZINC FINGER C3HC4.			
KW	ZINC-FINGER; DNA-BINDING; NUCLEAR PROTEIN.			
FT	ZN FING 19 58			
FT	DOMAIN 176 231			
FT	DOMAIN 285 348			
FT	DOMAIN 172 175			
SO	SEQUENCE 377 AA; 39145 MW; 716288 CN;			
Query Match	3.6%;	Score 151;	DB 6;	Length 377;

Best Local Similarity 24.7%; Pred. No. 1,41e-06;
Matches 24; Conservative 30; Mismatches 39; Indels 4; Gaps 4

RESULT	ID	7	STANDARD;	PRT;	1969 AA.
AC	DT	MYSA CAEEL			
AC	DT	P12844;			
DT	01-OCT-1989	(REL. 12, CREATED)			
DT	01-OCT-1989	(REL. 12, LAST SEQUENCE UPDATE)			
DT	01-JUN-1994	(REL. 29, LAST ANNOTATION UPDATE)			
DE	MYOSIN HEAVY CHAIN A (MHC A).				
GN	MYO-3.				
OC	CABERNADOTTIS ELEGANS.				
OC	ERKAROTIA; METRAGON; ACCELEMATES; NEMATODA; SECCENENTERA; RHADITITDA.				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-BRISTOL N2;				
RM	89178677				
RA	DIJB N.J., MARDYAMA I.N., KRAUSE M., KARN J.;				
RL	J. MOL. BIOL. 205:603-613(1988).				
CC	-1- FUNCTION: MUSCLE CONTRACTION.				
CC	-1- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2				
CC	HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)				
CC	AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).				
CC	-1- EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT MEROMYOSIN (LM)				
CC	AND 1 HEAVY MEROMYOSIN (HM). IT CAN LATTER BE SPLIT FURTHER INTO				
CC	2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED SUBFRAGMENT (S2).				
CC	-1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING				
CC	CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,				
CC	CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.				
CC	-1- SUBCELLULAR LOCATION: THICK FILAMENTS OF THE MYOFIBRILS.				
CC	-1- SIMILARITY: THE GLOBULAR HEAD AA SEQUENCE SHOWS A HIGH DEGREE OF				
CC	SIMILARITY WITH THE GLOBULAR HEAD SEQUENCES OF MUSCLE & NONMUSCLE				
CC	HEAVY CHAINS, BY CONTRAST THE ROD SEQUENCE IS LESS CONSERVED, BUT				
CC	THE PERIODICITIES OF HYDROPHOBIC & CHARGED RESIDUES, WHICH DICTATE				
CC	THE ALPHA-HELICAL COILED-COIL STRUCTURE, ARE CONSERVED.				
CC	-1- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY				
CC	ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.				
CC	-1- THERE ARE FOUR DIFFERENT MYOSIN HEAVY CHAINS IN C. ELEGANS.				
CC	-1- MHC A AND MHC B ARE FOUND EXCLUSIVELY IN THE BODY WALL MUSCLE.				
CC	THEY CO-ASSEMBLE INTO BODY WALL THICK FILAMENT.				
DR	EMBL; X08067; CEMY03.				
DR	PIR; S02771; S02771.				
DR	HSP7; P24733; 1SCX.				
RM	MYOSIN; MUSCLE PROTEIN; COILED COIL; THICK FILAMENT; ACTIN-BINDING;				
KW	ATP-BINDING; METHYLATION; ALKYLATION; HEPTAD REPEAT PATTERN;				
KW	MULTIGENE FAMILY.				
FT	DOMAIN	1	851	GLOBULAR HEAD (S1).	
FT	NP BIND	832	1969	RODLIKE TAIL (S2 AND LMM DOMAINS).	
FT		179	186	ATP.	
FT	DOMAIN	667	689	ACTIN-BINDING.	
FT	DOMAIN	770	784	ACTIN-BINDING.	

FT MOD_RES 130 130 METHYLATION (TRI-) (POTENTIAL).
 FT MOD_RES 707 707 ALKYLATION (SH-1) (POTENTIAL).
 FT MOD_RES 717 717 ALKYLATION (SH-2) (POTENTIAL).
 SQ SEQUENCE 1969 AA; 225509 MW; 1552438 CN;
 Query Match 3.6%; Score 151; DB 5; Length 1969;
 Best Local Similarity 21.2%; Pred. No. 1.41e-06;
 Matches 28; Conservative 49; Mismatches 47; Indels 8; Gaps 7;
 Db 985 dhnrlsrdmndeaavalknk-khgeenrklnedlqgeekvnhlekinklqeq 1043
 QY 270 EKKVSLQNESVEKKSISLHNOICFETELERQ-KEMLRNNSKILHLQWVIDSQAEK 328
 Db 1044 mdeleindrekeregkalekkrkveq-dlkvaq-enideitkq-kh-dv-etlikrke 1097
 QY 329 ILELDKEIRPRRQWNEADSMKSSVESLQNRVTELESVDKSGVARNRNGILESQLSRHD 388
 Db 1098 edlhnaklae 1109
 QY 389 QMLSVHDRLAD 400
 RESULT 8
 ID MYSN CHICK STANDARD; PRT; 1959 AA.
 AC P14105;
 DT 01-JAN-1990 (REL. 13, CREATED)
 DT 01-JAN-1990 (REL. 13, LAST SEQUENCE UPDATE)
 DT 01-JUN-1994 (REL. 29, LAST ANNOTATION UPDATE)
 DE MYOSIN HEAVY CHAIN, NONMUSCLE (NMHC).
 OS GALLUS GALLUS (CHICKEN).
 OC EDUAROTIA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; AVES; NEOGNATHAE;
 CC GALLIFORMES.
 CC [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=INTESTINAL EPITHELIUM;
 RM 90046668
 RA SHOHET R.V., CONTI M.A., KAWAMOTO S., PRESTON Y.A., BRILL D.A.,
 RA ADLSTEIN R.S.;
 RL PROC. NATL. ACAD. SCI. U.S.A. 86:7726-7730(1989).
 CC -1- FUNCTION: CELLULAR MYOSIN APPEARS TO PLAY A ROLE IN CYTOKINESIS,
 CELL SHAPE, AND SPECIALIZED FUNCTIONS SUCH AS SECRETION AND
 CAPING.
 CC -1- SUBUNIT: MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2 HEAVY
 CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC) AND 2
 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
 CC -1- EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT MEMOROSIN (LMM)
 AND 1 HEAVY MEMOROSIN (HMM). IT CAN LATTER BE SPLIT FURTHER INTO
 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED SUBFRAGMENT (S2).
 CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
 CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
 CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
 CC -1- SIMILARITY: BELONGS TO THE MYOSIN HEAVY CHAIN FAMILY. STRONGEST
 TO OTHER NONMUSCLE MYOSINS.
 CC EMBL; M26510; GGMYN.
 DR PIR; A33977; A33977.
 DR HSSP; P24733; 1SCM.
 KW MYOSIN; COILED COIL; ACTIN-BINDING; ALKYLATION; ATP-BINDING;
 HM7AD REPEAT PATTERN; MULTIGENE FAMILY.
 FT DOMAIN 1 835 GLOBULAR HEAD (S1).
 FT DOMAIN 836 1926 RODLIKE TAIL (S2 AND LMM DOMAINS).

FT NP BIND 174 181 ATP.
 FT DOMAIN 654 676 ACTIN-BINDING.
 FT MOD_RES 694 694 ALKYLATION (SH-1) (POTENTIAL).
 FT MOD_RES 704 704 ALKYLATION (SH-2) (POTENTIAL).
 SQ SEQUENCE 1959 AA; 226502 MW; 15394260 CN;
 Query Match 3.5%; Score 145; DB 5; Length 1959;
 Best Local Similarity 25.2%; Pred. No. 9.59e-06;
 Matches 34; Conservative 34; Mismatches 57; Indels 10; Gaps 8;
 Db 1460 yaeedraaeareketakalaleaeleqaeleerunkqtr-t-emedlmskcdv 1516
 QY 297 FELIERQKEMLRNNSKILHLQWVIDSQAEKLELDKEIRPRRQWNEADSMKSSVESL 356
 Db 1517 gksvhelekakrallqeqveemktql-e-el--edelqatadaklr-levngam-kagidr 1571
 QY 357 QNRVTELESVDKS-AGGVARNRNGILESQLSRHDMVSDIRLADMDLGFVLETRSYN- 414
 Db 1572 dlgrdeqneekrkq 1586
 QY 415 GVLIWKIRIDYKRRQ 429
 RESULT 9
 ID MYSN HUMAN STANDARD; PRT; 1961 AA.
 AC P35579;
 DT 01-JUN-1994 (REL. 29, CREATED)
 DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)
 DT 01-JUN-1994 (REL. 29, LAST ANNOTATION UPDATE)
 DE MYOSIN HEAVY CHAIN, NONMUSCLE TYPE A (CELLULAR MYOSIN HEAVY CHAIN,
 DE TYPE A) (NMHC-A).
 OS MYH9.
 CC HOMO SAPIENS (HUMAN).
 CC EDUAROTIA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 CC EUTHERIA; PRIMATES.
 CC [1]
 RP SEQUENCE OF 1-1337 FROM N.A.
 RM 92003925
 RA TOOTHAKER L.E., GONZALEZ D.A., TUNG N., LEMONS R.S., LE BEAU M.M.,
 RA ARNAOUT M.A., CLAYTON L.K., TENEN D.G.;
 RL BLOOD 78:1826-1833(1991).
 CC [2]
 RP SEQUENCE OF 1-715 FROM N.A.
 RM 91316803
 RA SIMONS M., WANG M., MCBRIDE O.W., KAWAMOTO S., YAMAKAWA K.,
 RA GDULA D., ADLSTEIN R.S., WEIR L.;
 RL CIRC. RES. 69:530-539(1991).
 CC [3]
 RP SEQUENCE OF 715-1961 FROM N.A.
 RM 90138958
 RA SAEZ C.G., MYERS J.C., SHOWS T.B., LEINWAND L.A.;
 RL PROC. NATL. ACAD. SCI. U.S.A. 87:1164-1168(1990).
 CC -1- FUNCTION: CELLULAR MYOSIN APPEARS TO PLAY A ROLE IN CYTOKINESIS,
 CELL SHAPE, AND SPECIALIZED FUNCTIONS SUCH AS SECRETION AND
 CAPING.
 CC -1- SUBUNIT: MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2 HEAVY
 CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC) AND 2
 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
 CC -1- EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT MEMOROSIN (LMM)
 AND 1 HEAVY MEMOROSIN (HMM). IT CAN LATTER BE SPLIT FURTHER INTO

ID	RESULT	11	STANDARD;	PRT; 1978 AA.
AC	MYSG CHICK			
DT	P10587;			
DT	01-JUL-1989 (REL. 11, CREATED)			
DT	01-JAN-1990 (REL. 13, LAST SEQUENCE UPDATE)			
DE	01-JUN-1994 (REL. 29, LAST ANNOTATION UPDATE)			
DE	MYOSIN HEAVY CHAIN, GIZZARD SMOOTH MUSCLE.			
OS	GALLUS GALLUS (CHICKEN).			
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; AVES; NEOGATIAE; GALLIFORMES.			
RN	[1]			
RN	SEQUENCE FROM N.A.			
RM	88116918			
RA	YAMAGISHIMA M., HAMADA Y., KATSURAGAWA Y., IMAMURA M., MIKAWA T., MASAKI T.;			
RL	J. MOL. BIOL. 198:143-157(1987).			
RP	[2]			
RP	REVISIONS.			
RA	MASAKI T.;			
RL	SUBMITTED (FEB-1989) TO EMBL/GENBANK/DDBJ DATA BANKS.			
RP	[3]			
RP	SEQUENCE OF 1-203.			
RM	88032919			
RA	MAITA T., ONISHI H., YAJIMA E., MATSUDA G.;			
RL	J. BIOCHEM. 102:133-145(1987).			
CC	-1- FUNCTION: MUSCLE CONTRACTION.			
CC	-1- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2 HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC) AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).			
CC	-1- EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT MEROMYOSIN (LMH) AND 1 HEAVY MEROMYOSIN (HMH). IT CAN LATER BE SPLIT FURTHER INTO 2 GLOBALAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED SUBFRAGMENT (S2).			
CC	-1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES. CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.			
CC	-1- SUBCELLULAR LOCATION: THICK FILAMENTS OF THE MYOFIBRILS.			
CC	-1- SIMILARITY: THE PERIODICITIES OF HYDROPHOBIC AND CHARGED RESIDUES, WHICH DICTATE THE ALPHA-HELICAL COILED-COIL STRUCTURE ARE CONSERVED.			
CC	-1- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.			
DR	EMBL; X06346; GENMFGSM.			
DR	PIR; S03166; S03166.			
DR	HSSP; P02259; IGHC.			
KW	MYOSIN; MUSCLE PROTEIN; COILED COIL; THICK FILAMENT; ACTIN-BINDING; APP-BINDING; METHYLATION; ALKYLATION; HEPTAD REPEAT PATTERN; MULTIGENE FAMILY.			
FT	INIT MET	0		
FT	MOD RES	1		
FT	DOMAIN	1	848	BLOCKED.
FT	DOMAIN	849	1978	GLOBULAR HEAD (S1).
FT	NP BIND	176	183	RODLIKE TAIL (S2 AND LMH DOMAINS).
FT	DOMAIN	666	688	APP.
FT	DOMAIN	767	781	ACTION-BINDING.
FT	MOD RES	127	127	ACTIN-BINDING.
				METHYLATION (TRI-) (POTENTIAL).

Query Match	3.2%;	Score 135;	DB 5;	Length 1978;
Best Local Similarity	23.1%;	Pred. No. 2.14e-04;		
Matches	36;	Conservative	44;	Mismatches 69;
			Indels	7;
			Gaps	7;

Dh 1037 heamsellevrllkkeeekargetelekrrkllegesadhbqaelgaqi;aelkaqlakkeee 1096
 Qy 250 HEASVVOHVNIITE-WS-NLEKRVSLIOMNSVEKNKSIOGLNHOICSEIEIEROKEM 307
 Dh 1097 lgaalarladeresqkmalkkireleeshisadiqedlesekaarakaek-qkrilseelea 1155
 Qy 308 LRNNESKIILHLORVIDSQAEEKELKDKEIRPRONNELEADSKSVESLONR-VTE-LTS 365
 Dh 1156 1-kteledtldtarqgelrakregevtvfkralee 1190
 Qy 366 VDKSAGQAVANTGELLESQL-SRDHDOMSHVDRLAD 400

RESULT	12	STANDARD;	PRT;	245	AA.
ID	TPWZ_RAT				
AC	P18344;				
DT	01-NOV-1990 (REL. 16, CREATED)				
DT	01-NOV-1990 (REL. 16, LAST SEQUENCE UPDATE)				
DT	01-NOV-1990 (REL. 16, LAST ANNOTATION UPDATE)				
DE	TROPOMYOSIN ALPHA CHAIN, BRAIN-3 (TMBR-3).				
GN	ALPHA-TM.				
OS	RATTUS NORVEGICUS (RAT).				
OC	EMBRYONIC, METRIZOL; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;				
OC	ETHEREALIA; RODENTIA.				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=BRAIN;				
RM	90205854				
RA	LEES-MILLER J.P., GOODWIN L.O., HELFMAN D.M.;				
RL	MOL. CELL. BIOL. 10:1729-1742(1990).				
CC	-1- FUNCTION: THE FUNCTION OF TROPOMYOSIN IN SMOOTH MUSCLE AND NON-				
CC	MUSCLE CELLS IS NOT CLEAR.				
CC	-1- SUBUNIT: DIMER OF AN ALPHA AND A BETA CHAIN.				
CC	-1- DOMAIN: THE MOLECULE IS IN A COILED COIL STRUCTURE. THE SEQUENCE				
CC	EXHIBITS A PROMINENT SEVEN-RESIDUES PERIODICITY.				
CC	-1- TISSUE SPECIFICITY: BRAIN.				
CC	-1- ALTERNATIVE PRODUCTS: THE TROPOMYOSIN ALPHA-TM GENE PRODUCES, BY				
CC	ALTERNATIVE SPLICING, THE FOLLOWING PROTEINS: STRIATED MUSCLE TM,				
CC	SMOOTH MUSCLE TM, BRAIN TMBR-1, TMBR-2, TMBR-3, FIBROBLAST TM-2,				
CC	TM-3, TM-5A, AND TM-5B.				
CC	EMBL: M34136; RNTMBR3A.				
DR	PIR: C34787; C34787.				
DR	PROSITE; P500326; TROPOMYOSIN.				
DR	BRAIN COILED COIL; REPEAT; ALTERNATIVE SPLICING; MULTIGENE FAMILY.				
KW	SEQUENCE 245 AA; 28343 MW; 256960 CN;				
SO					
	Query Match	3.1%;	Score 128;	DB 7;	Length 245;
	Best Local Similarity	23.5%;	Pred. No. 1,75e-03;		
	Matches 20;	Conservative 28;	Mismatches 34;	Indels 3;	Gaps 3;
DB	31 grelid-qerk retcadvaal nr q veel drqer at alq beek adeeer 89				

DE TROPOMYOSIN, MUSCLE.
OS TRICHOSTRONGYLUS COLDERIFORMIS.
OC EUDAROTIA; METAZOA; ACCELMATES; NEMATODA; SECHERNENTEA; STRONGYLIDA.
RN [1]
RP SEQUENCE FROM N.A.
RM 90114326
RA FRENKEL M.J., SAVIN K.W., BAKER R.E., WARD C.W.;
RL MOL. BIOCHEM. PARASITOL. 37:191-200(1989).
RN [2]
RP PARTIAL SEQUENCE.
RM 89339870
RA O'DONNELL I.J., DINEEN J.K., WAGLAND B., LETHO S., WERKMEISTER J.A.,
WARD C.W.;
RL INT. J. PARASITOL. 19:327-335(1989).
CC -1- FUNCTION: TROPOMYOSIN, IN ASSOCIATION WITH THE TROPONIN COMPLEX,
PLAYS A CENTRAL ROLE IN THE CALCIUM DEPENDENT REGULATION OF
STRIATED MUSCLE CONTRACTION.
CC -1- SUBUNIT: HOMODIMER.
CC -1- DOMAIN: THE MOLECULE IS IN A COILED COIL STRUCTURE. THE SEQUENCE
EXHIBITS A PROMINENT SEVEN-RESIDUES PERIODICITY.
CC -1- DEVELOPMENTAL STAGE: PRESENT IN L3 (THIRD STAGE), L4 AND ADULT
WORMS.
DR EMBL; J04669; TCTROP.
DR HSP; P07751; 2SPC.
DR PROSITE; P500325; COFILIN_TROPOMYOSIN.
DR PROSITE; P500326; TROPOMYOSIN.
KM MUSCLE PROTEIN; COILED COIL; REPEAT.
FT CONFLICT 277 277 T -> R (IN REF. 2).
SQ SEQUENCE 284 AA; 33050 MW; 355363 CN;

Query Match 3.1%; Score 131; DB 7; Length 284;
Best Local Similarity 20.5%; Pred. No. 7.18e-04;
Matches 25; Conservative 49; Mismatches 42; Indels 6; Gaps 6;

Db 85 vaalnrmllleeleraerlkatekleathndv-eservkv-mengsfde-era 141
Qy 287 IQSLHNOICSEIEIRKOKELRNNSKILHROVIDSOAEKKEKIDKIRPFRONWEA 346
Db 142 ntiagkkae-miae-ea-drydevarklamveadleraeaeagenkiveleelr 198
Qy 347 DSMSSVESLONRYTELESVDKSGAVARTGILESQLSRHOMLSVHDRIADMDLGFQ 406
Db 199 vv 200
Qy 407 VL 408

Search completed: Tue Mar 19 09:51:46 1996
Job time : 35 secs.

WVPSRLH (TM)

Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by Intelligentics, Inc.

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm
Run on: Tue Mar 19 09:55:10 1996; Maspar time 1.86 Seconds
190.366 Million cell updates/sec
Tabular output not generated.

Title: >US-08-404-862-3
Description: (1-49) from US08404832.pep
Perfect Score: 377
Sequence: 1 DKYKCEHVLICSPKOTECGHCRCSCMAALLSSSPKCTACQESIVK 49

Scoring table: PAM 150
Gap 11

Searched: 62395 seqs, 7230759 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: a-geneeq20
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11 12:part12

Statistics: Mean 22.842; Variance 90.056; scale 0.254

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	DB	ID	Description	Pred. No.
1	107	28.4	475	3	R15148	Ro/SSA autointgen.	1.08e-02
2	83	22.0	69	11	R60623	Herpes simplex virus	1.49e+00
3	82	21.8	365	1	P91461	Rpt-1 clone cDNA inse	1.82e+00
4	76	20.2	560	5	R27535	myl protein.	5.90e+00
5	76	20.2	797	5	R27533	myl/RAR-alpha fused.	5.90e+00
6	74	19.6	165	2	R10533	Prod. of pMG4B12 used.	8.69e+00
7	73	19.4	67	11	R60622	Pseudorabies virus ea	1.05e+01
8	73	19.4	410	11	R60621	Pseudorabies virus ea	1.05e+01
9	72	19.1	574	12	R66033	Human ARD 1.	1.28e+01
10	71	18.8	574	12	R66034	Rat ARD 1.	1.54e+01
11	69	18.3	124	12	R63059	Rat PLIA2 type I.	2.26e+01
12	69	18.3	146	12	R63052	HLA2-8.	2.26e+01
13	69	18.3	185	10	P53530	Partial Thyroid hormo	2.26e+01
14	69	18.3	555	1	R05434	Sequence for lucifera	2.26e+01
15	69	18.3	555	2	R11108	Luciferase deduced fr	2.26e+01
16	68	18.0	118	1	P90076	Phospholipase A2-type	2.72e+01
17	67	17.8	104	6	R30630	Polypeptide coded by	3.28e+01
18	67	17.8	104	6	R30631	Polypeptide coded by	3.28e+01
19	67	17.8	2783	5	R23963	APP-1 (Ala 2460 Val).	3.28e+01
20	67	17.8	2783	5	R23962	APP-1.	3.28e+01

21	66	17.5	208 11	R60055	Dirofilaria immitis p	3.96e+01
22	65	17.2	51 12	R69821	OMTKY3(6-56) (A15R-T17	4.77e+01
23	64	17.0	56 3	R14310	Tomato metallothionein	5.73e+01
24	64	17.0	77 6	R34939	p270 polypeptide.	5.73e+01
25	64	17.0	81 2	R10311	Ovary tissue transcri	5.73e+01
26	64	17.0	149 7	R39206	Sequence of a peptide	5.73e+01
27	63	16.7	2458 2	R04031	Full length T4 encode	6.89e+01
28	62	16.4	2458 2	R07640	Deduced protein sequ	6.89e+01
29	62	16.4	103 12	R70984	Component B protein.	8.27e+01
30	62	16.4	129 2	R12113	Lysosyme substrate fo	8.27e+01
31	62	16.4	146 9	R53348	Chicken (egg white) 1	8.27e+01
32	62	16.4	154 3	P50461	Sequence of the anti	8.27e+01
33	62	16.4	154 2	P82174	Peptide encoded by HB	8.27e+01
34	62	16.4	291 1	R05596	Somatomedin carrier p	8.27e+01
35	62	16.4	442 11	R60536	Mouse osteoblast-spec	8.27e+01
36	62	16.4	466 1	R07447	Human laminin B1 cha	8.27e+01
37	62	16.4	635 4	R23970	MPLV env protein with	8.27e+01
38	62	16.4	873 8	R44735	apo-E lipoprotein rec	8.27e+01
39	61	16.2	67 11	R60624	Varicella zoster viru	9.91e+01
40	61	16.2	622 12	R68840	Plasmodium falciparum	9.91e+01
41	61	16.2	622 1	P91632	Rhoptry membrane anti	9.91e+01
42	61	16.2	706 12	R68743	BCL-6 zinc finger pro	9.91e+01
43	61	16.2	718 1	R05936	Secreted Grp78 subun	9.91e+01
44	61	16.2	1141 12	R66391	Human SRBP-2.	9.91e+01
45	61	16.2	1764 1	P91672	Primary amino acid se	9.91e+01

ALIGNMENTS

RESULT 1
ID R15148 standard; Protein; 475 AA.

AC R15148;
DT 14-FEB-1992 (first entry)
DE Ro/SSA autoantigen.
KW Autoantibody; autoantigen; SLE; systemic lupus erythematosus.
OS Homo sapiens.
PN W091171-A.
PD 14-NOV-1991.
PF 07-MAY-1991; D03139.
PR 07-MAY-1990; US-520270.
PA (OKLA-) OKLAHOMA MED RES FO.
PI Frank MB, Itoh K,
DR WPI; 91-353712/48.
DR N-PSDB; Q14798.
PT DNA encoding an Ro-SSA autoantigen - useful for diagnosing
PT auto-immune disorders or presence of auto-antibodies
PS Disclosure; Fig 2; 41pp; English.
CC A cDNA library (from human thymus mRNA) in lambda gtl1 was screened
CC with serum from a patient having systemic lupus erythematosus. Two
CC clones were reactive with sera (from a panel of lupus patients)
CC which contd. autoantibodies against 52 kd protein.
CC Both the cDNA and the protein expressed from it, or portions of it,
CC are useful as diagnostic agents in the identification of patients
CC having autoantibodies and in the identification and analysis of
CC the structural and functional properties of the autoantigen and for
CC application in immunotherapeutic regimens.
SQ Sequence 475 AA;

Query Match 28.4%; Score 107; DB 3; Length 475;
Best Local Similarity 27.3%; Pred. No. 1.08e-02;

Matches	12;	Conservative	13;	Mismatches	19;	Indels	0;	Gaps	0;
Db	16	cpicldpfpvpsiechscfscgcicqvgkgygsvoverqf1	59						
Qy	5	CEKCHLVICSPKQTEGCHRCESOMALLSSSPKCTAOQESIV	48						

RESULT 2
ID R60623 standard; Protein; 69 AA.
AC R60623;
DT 15-MAY-1995 (first entry)
DE Herpes simplex virus ICPO protein domain homologous to part of EPO.
KW Pseudotables virus; PRV; LTV; large latency transcript;
KW attenuated virus; vaccine; early protein 0; EPO; HSV-1 ICPO;
KW protecting animals; deletion mutants; swine; cysteine-rich;
KW zinc finger motif.
OS Herpes simplex virus 1.
PI Key
FH Location/Qualifiers
FT Misc difference 6
FT /note= "cysteine residue that is part of the zinc
FT finger motif"
FT Misc difference 30
FT /note= "cysteine residue that is part of the zinc
FT finger motif"
FT Misc difference 41
FT /note= "cysteine residue that is part of the zinc
FT finger motif"
FT Misc difference 44
FT /note= "cysteine residue that is part of the zinc
FT finger motif"
FT US3532596-A.
PN 04-OCT-1994.
PD 11-SEP-1992; 945283.
PF 11-SEP-1992; US-945283.
PR (USDA) US SEC OF AGRIC.
PA Cheung AK, Wesley RD;
PI WPI; 94-316187/39.
DR New pseudotables virus mutants for use in vaccine - having a
PT deletion and/or insertion in the early protein 0 gene or large
PT latency transcript gene
PS Disclosure; Column 45-47; 43pp; English.
CC R60622-24 show cysteine-rich zinc finger domains that is conserved among
CC herpesviruses. This peptide is the homologous region from the Herpes
CC simplex virus 1. This is the most important domain of the EPO gene. The
CC DNA encoding this cysteine-rich domain (see Q73501) also encodes an amino
CC acid sequence specific to the LIT gene in the opposite orientation. Thus
CC deletion in the EPO gene other than the first 200 bp of EPO will
CC automatically delete some of the DNA sequences encoding the LTV. EPO is
CC nonessential for replication, LIT (large latency transcript) is the only
CC gene expressed during PRV latency, and the IE180 gene is absolutely
CC necessary for PRV replication. However there are 2 copies of IE180 in the
CC genome. It is expected that PRV lacking one of the IE180 copies is
CC viable. Deletions in the non-overlapping regions of these 3 genes will
CC generate single deletion routants, while deletions in overlapping regions

CC will generate double deletion mutants. The invention is concerned with
 CC the construction of attenuated viruses which have a reduced ability to
 CC reactivate from latency. This can be achieved by functionally disabling
 CC the expression of the EPO gene, or by disrupting the synthesis of the
 CC LIT, or both. (See also Q73500-1 and R60620-24)
 SQ Sequence 69 AA;

Query Match 22.0%; Score 83; DB 11; Length 69;
 Best Local Similarity 34.5%; Pred. No. 1.49e+00;
 Matches 10; Conservative 5; Mismatches 13; Indels 1; Gaps 1;

Db 24 cmhfcjpcmkvmqint-cplonaklv 51
 QY 20 CGHRFCSCMAALITSSSPKCTACQESIV 48

RESULT 3
 ID P91461 standard; peptide; 365 AA.
 AC P91461;
 DT 10-MAR-1993 (revised)
 DT 09-JAN-1990 (first entry)
 DE Rpt-1 clone cDNA insert.
 KW Regulatory protein T lymphocyte-1; IL-2 receptor;
 KW human lymphotropic retroviruses; metal fingers.
 OS Homo sapiens.
 FN W08907652-A.
 PD 24-AUG-1989.
 PF 10-FEB-1989; US-154758.
 PR 11-FEB-1988; US-154758.
 PA (DANA) Dana-Farber Cancer Institute.
 PI Cantor HI, Patarca R, Freeman GJ;
 DR WPI; 89-263719/36.
 PT Regulatory protein T lymphocyte-1 gene and corresp. protein - increase
 PT gene expression of interleukin-2 receptor and of human lymphotropic
 PT retroviruses.
 PS Claim 20; fig 2; 58pp; English.
 CC cDNA was prep'd. from T cells (fibroblast tumour) and 2PK3 (B cell
 CC lymphoma). Host was pref. E. coli NRRL B-18297. Cysteine residues at
 CC positions 15, 18, 35, 38, 55, 58, 115 and 118, and histidine residues at
 CC positions 124 and 127 may be involved in metal finger formation.
 CC Positions are given from the methionine (a.a 13 in the sequence given).
 CC See also N90702.
 SQ Sequence 365 AA;

Query Match 21.8%; Score 82; DB 1; Length 365;
 Best Local Similarity 34.6%; Pred. No. 1.82e+00;
 Matches 9; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

Db 27 cplclellkpsadchsfrcacit 52
 QY 5 CECHIVLCSPKQTECGHRFCSCMA 30

RESULT 4
 ID R27535 standard; Protein; 560 AA.
 AC R27535;
 DT 09-MAR-1993 (first entry)
 DE myl protein.
 KW Retinoic acid receptor; RAR-alpha; myl; acute promyelocytic leukemia;
 KW APL; translocation; chromosome 17; chromosome 15; PCR; primer;

KW [t(15;17)(q21;q11-22)]; breakpoint; polymerase chain reaction.

OS Synthetic.
 PN W09216660-A.
 PD 01-OCT-1992.
 PF 23-MAR-1992; D02320.
 PR 22-MAR-1991; US-673838.
 FR 22-MAR-1991; US-675084.
 PA (SLOK) SLOAN KETTERING INST CANCER.
 PI Dmitrovsky E, Evans RM, Frankel S, Kazizuka A, Miller WH;
 PI Warrell RJ;
 DR WPI; 92-349240/42.
 DR N-PSDB; Q29354.

PT Marker for acute promyelocytic leukaemia and other neoplasias -
 PT comprising nucleic acid and encoded abnormal retinoic acid
 PT receptor-alpha receptor
 PS Disclosure; Page 48-51; 84pp; English.
 CC The sequence given represents the amino acid sequence of one variation
 CC of the myl protein. The gene encoding this protein is disrupted in a
 CC translocation of a portion of the long arm of chromosome 17 onto the
 CC long arm of chromosome 15 [t(15;17)(q21;q11-22)]. This causes a fusion
 CC between retinoic acid receptor (RAR)-alpha and myl which is
 CC characteristic of acute promyelocytic leukemia (APL). The breakpoint
 CC region has been cloned and it has been shown that DNA rearrangements
 CC are clustered in the region of the first intron of RAR-alpha. This
 CC sequence was isolated by polymerase chain reaction (PCR). The primers
 CC used for amplification of this sequence can also be used to amplify the
 CC translocated region.
 SQ Sequence 560 AA;

Query Match 20.2%; Score 76; DB 5; Length 560;
 Best Local Similarity 34.5%; Pred. No. 5.90e+00;
 Matches 10; Conservative 6; Mismatches 11; Indels 2; Gaps 2;

Db 56 rcqgqaekc-pklpnlhtlcagclea 83
 QY 4 KCERCHL-VLCSPKQTECGHRFCSCMA 31

RESULT 5
 ID R27533 standard; Protein; 797 AA.
 AC R27533;
 DT 09-MAR-1993 (first entry)
 DE myl/RAR-alpha fusion.
 KW Retinoic acid receptor; RAR-alpha; myl; acute promyelocytic leukemia;
 KW APL; translocation; chromosome 17; chromosome 15; PCR; primer;
 KW [t(15;17)(q21;q11-22)]; breakpoint; polymerase chain reaction.
 OS Synthetic.
 PN W09216660-A.
 PD 01-OCT-1992.
 PF 23-MAR-1992; D02320.
 PR 22-MAR-1991; US-673838.
 PR 22-MAR-1991; US-675084.
 PA (SLOK) SLOAN KETTERING INST CANCER.
 PI Dmitrovsky E, Evans RM, Frankel S, Kazizuka A, Miller WH;
 PI Warrell RJ;
 DR WPI; 92-349240/42.
 DR N-PSDB; Q29334.
 PT Marker for acute promyelocytic leukaemia and other neoplasias -
 PT comprising nucleic acid and encoded abnormal retinoic acid
 PT receptor-alpha receptor

PS Disclosure; Fig 4; 84pp; English.
 CC The sequence given shows a fusion between retinoic acid receptor
 CC (RAR)-alpha and myl which is characteristic of acute promyelocytic
 CC leukemia (APL). This is caused by a translocation of a portion of
 CC the long arm of chromosome 17 onto the long arm of chromosome 15
 CC [t(15;17)(q21;q11-22)]. The breakpoint region has been cloned and it
 CC has been shown that DNA rearrangements are clustered in the region
 CC of the first intron of RAR-alpha. This sequence was isolated by
 CC polymerase chain reaction (PCR) using primers which correspond to
 CC sequences both 5' and 3' to the breakpoint region.
 SQ Sequence 797 AA;

Query Match 20.2%; Score 76; DB 5; Length 797;
 Best Local Similarity 34.5%; Pred. No. 5.90e+00;
 Matches 10; Conservative 6; Mismatches 11; Indels 2; Gaps 2;

Db 56 rcgqgaakc-pkllpchlntlcscglea 83
 Qy 4 KCEKCHL-VLCSPKTECGHRCSCMAA 31

RESULT 6
 ID R10533 standard; Protein; 165 AA.
 AC R10533;
 DT 12-APR-1991 (first entry)
 DE Prod. of pMG4B12 used to isolate style-stigma specific gene STG4B12.
 KW Hybrid vigour; pollination.
 OS Nicotiana tabacum "Petit Havana" SRI.
 PN EP-412006-A.
 PD 06-FEB-1991.
 PF 31-JUL-1990; 402196.
 PR 04-AUG-1989; EP-402224.
 PR 31-JUL-1990; EP-402196.
 PA (PLAN-) PLANT GENETIC SYST.
 PI De Greef W, Van Emmelo J, De Oliveira DE, De Souza MH;
 PI Van Montagu M;
 DR N-PSDB; 010376.
 DR WPI; 91-038878/06.
 PT Modified plant with transformed nuclear genome - obtd. by using
 PT foreign DNA encoding prod. which disrupts metabolism, functioning
 PT and/or development of plant.
 PS Disclosure; Fig 2b; 42pp; English.
 CC The clone was produced by subcloning a style-stigma specific cDNA,
 CC 4B12, obtd. from Prof. Goldberg of UCLA, into pGEM1, to produce
 CC pMG4B12. The clone was found to be stigma-style specific in
 CC Northern analysis. A probe from the clone was used to isolate
 CC the corresp. genomic sequence which is specifically expressed in
 CC style-stigma tissues of the female organ of tobacco. The corresp.
 CC clone, pSTG4B12, contains STMG4B12, an "STMG-type" gene. This and
 CC other similar genes can be used to produce new hybrid plants or
 CC seeds having a combination of desirable traits, and showing hybrid
 CC vigour. A female-sterile, male-fertile plant can be produced which
 CC favours cross-pollination.
 CC See also R10533-35.
 SQ Sequence 165 AA;

Query Match 19.6%; Score 74; DB 2; Length 165;
 Best Local Similarity 47.8%; Pred. No. 8.69e+00;
 Matches 11; Conservative 3; Mismatches 7; Indels 2; Gaps 2;

Db 143 ncgkcnv-cspgqk-csfqcd 163
 Qy 4 KCEKCHLVCSPKTECGHRCFCE 26

RESULT 7

ID R60622 standard; Protein; 67 AA.
 AC R60622;
 DT 15-MAY-1995 (first entry)
 DE Pseudorabies virus early polypeptide 0 homologous protein domain.
 DE Pseudorabies virus; PRV; LIV; large latency transcript;
 KW Pseudorabies virus; vaccine; early protein 0; EPO; HSV-1 ICP0;
 KW attenuated virus; deletion mutants; swine; cysteine-rich;
 KW protecting animals; deletion mutants; swine; cysteine-rich;
 OS Pseudorabies virus.
 PN Key
 PI Key

FT Misc difference 6 Location/Qualifiers
 FT /note= "cysteine residue that is part of the zinc
 FT finger motif"
 FT Misc difference 9
 FT /note= "cysteine residue that is part of the zinc
 FT finger motif"
 FT Misc difference 27
 FT /note= "cysteine residue that is part of the zinc
 FT finger motif"
 FT Misc difference 30
 FT /note= "cysteine residue that is part of the zinc
 FT finger motif"
 FT Misc difference 41
 FT /note= "cysteine residue that is part of the zinc
 FT finger motif"
 FT Misc difference 44
 FT /note= "cysteine residue that is part of the zinc
 FT finger motif"
 FT US352596-A.
 PN 04-OCT-1994.
 PD 11-SEP-1992; 945283.
 PF 11-SEP-1992; US-945283.
 PR (USDA) US SEC OF AGRIC.
 PA Cheung AK, Wesley RD;
 PI WPI; 94-316187/39.
 DR New pseudorabies virus mutants for use in vaccine - having a
 PI deletion and/or insertion in the early protein 0 gene or large
 PI latency transcript gene
 PS Disclosure; Column 45-46; 43pp; English.
 CC R60622-24 show cysteine-rich zinc finger domains that is conserved among
 CC herpesviruses. This is the most important domain of the EPO gene. The DNA
 CC encoding this cysteine-rich domain (see Q73501) also encodes an amino
 CC acid sequence specific to the LIT gene in the opposite orientation. Thus
 CC deletion in the EPO gene other than the first 200 bp of EPO will
 CC automatically delete some of the DNA sequences encoding the LIT. EPO is
 CC nonessential for replication, LIT (large latency transcript) is the only
 CC gene expressed during PRV latency, and the IE180 gene is absolutely
 CC necessary for PRV replication. However there are 2 copies of IE180 in the
 CC genome. It is expected that PRV lacking one of the IE180 copies is
 CC viable. Deletions in the non-overlapping regions of these 3 genes will
 CC generate single deletion mutants, while deletions in overlapping regions
 CC will generate double deletion mutants. The invention is concerned with
 CC the construction of attenuated viruses which have a reduced ability to
 CC reactivate from latency. This can be achieved by functionally disabling

Query Match	19.4%;	Score 73;	DB 11;	Length 67;
Best Local Similarity	36.2%;	Pred. No. 1.05e+01;		
Matches 17; Conservative		7;	Mismatches 20;	Indels 3; Gaps 2.

RESULT	8
ID	R60621; standard; Protein; 410 AA.
AC	R60621;
DT	15-MAY-1995 (first entry)
DE	Pseudorabies virus early polypeptide 0.
KW	Pseudorabies virus; PRV; LIV; large latency transcript;
KM	attenuated virus; vaccine; early protein 0;
KN	EP0; HSV-1 ICP0;
KS	protecting animals; deletion mutants; swine.

	Key	Location/Qualifiers
PH	domain	40...100
FT	/note=	"cysteine rich zinc finger domain"
FT	domain	46...49
FT	/label=	zinc_finger_motif
FT	/note=	"cysteine rich"
FT	domain	67...70
FT	/label=	zinc_finger_motif
FT	/note=	"cysteine rich"
FT	domain	81...84
FT	/label=	zinc_finger_motif
FT	/note=	"cysteine rich"
PN	US3353596-A.	
PD	04-OCT-1994.	
PF	11-SEP-1992;	945283.
PR	11-SEP-1992;	US-945283.
PA	(USDA)	US SEC OF AGRIC.
PI	Cheung AK,	Wesley RD;
DR	WFI;	94-316187/39.
DR	N-PDSB;	Q73501.
PT	New pseudocapsids	virus mutants for use in vaccine - having a
PT	latency and/or	insertion in the early protein 0 gene or large
PT	latency transcript	gene
PS	Disclosure;	Column 39-44; 43pp; English.
CC	R60621	shows the early polypeptide 0 (EP0) encoded by Q73501. The
CC	most	important of the EP0 gene is likely to be the cysteine-rich zinc
CC	finger	domain from amino acids 40 to 100, since this region is conserved
CC	among	other herpesviruses (eg. see R60622-24). The DNA encoding this
CC	cysteine-rich	domain also encodes an amino acid sequence specific to the
CC	LIT	gene in the opposite orientation. Thus deletion in the EP0 gene
CC	other	than the first 200 bp of EP0 will automatically delete some of the
CC	DNA	sequences encoding the LIT. EP0 is nonessential for replication, LIT
CC	(large	latency transcript) is the only gene expressed during PRV latency
CC	and	the IE180 gene is absolutely necessary for PRV replication. However
CC	there	are 2 copies of IE180 in the genome. It is expected that PRV
CC	lacking	one of the IE180 copies is viable. Deletions in the
CC	non-overlapping	regions of these 3 genes will generate single deletion
CC	routeants,	while deletions in overlapping regions will generate double

Query Match 19.4%; Score 73; DB 11; Length 410;
Best Local Similarity 36.2%; Pred. No. 1.05e+01;
Matches 17; Conservative 7; Mismatches 20; Indels 3; Gaps 2;

RESULT	9
ID	R66033 standard; Protein; 574 AA.
AC	R66033;
DT	05-JUN-1995 (first entry)
DE	Human ARD 1.
KW	ARD-ribosylation factor; ARF domain; ARD 1
OS	Homo sapiens.
PN	W039424283.A.
PD	27-OCT-1994.

PR	16-APR-1993; US-0492352.
PR	19-APR-1993; US-049473.
PA	(USSH) US DEPT HEALTH & HUMAN SERVICES.
PI	Mishima K, Moss J, Nightingale M, Tsuchiya M, Wetj 94-341862/42.
DR	
PT	N-PDSDB; Q79326.
PT	GTP-binding protein ARD1 with ADP-ribosylation factor domain - useful as biochemical and diagnostic reagent
PS	Claim 5; Page 24-27; 52pp; English.
CC	The novel ARD 1 protein includes an 18 kDa region that exhibits significant homology to known ADP-ribosylation factors (ARFs) but lacks a 15 AA domain previously thought necessary for ARF stimulation. The 18 kDa region is called the functional domain and it is at the carboxy terminus. It is called ARD 1 for ARF domain. A recombinant truncated species contg. only the ARF domain following expression activated cholera toxin ADP-ribosyltransferase in a GTP- dependent manner, consistent with the conclusion that 15 AA adjacent to the amino terminus of ARF proteins are not required for toxin activation. cDNA was isolated from a human HL-60 lambda library screened with ARF 2B cDNA and a mixture of oligos denoted XARC. Clone no. 76 was positive with the ARF 2B cDNA and negative with oligos specific for ARFs 1-6. The insert was sequenced. The insert (bps 706-2365 Q79326) included an ORF (1207-1722) encoding an ARF domain of 172 AAs. Oligo JIR was used to screen a human fetal brain cDNA lambda ZAP library. Clones were found contg. nucleotides 7-1826 and 726-3225 of Q79326. In this sequence, about 1200 nts preceded the ARF region without a stop codon in the same ORF. To further characterise the 5' terminus of this cDNA, 5'-RACE was carried out with the poly A+ RNA from IMR-32 human neuroblastoma cells.
SQ	Sequence 574 AA;
Query Match	19.1%; Score 72; DB 12; Length 574;
Best Local Similarity	33.3%; Pred. No. 1.28e+01;

QY 13 CSPKOTEGHRCESCMALLSSS 36

RESULT 15
ID R1108 standard; Protein; 555 AA.
AC R1108;
DE 23-MAY-1991 (first entry)
DE Luciferase deduced from cDNA.
KW Vargula hilgendorffii.
OS Vargula hilgendorffii.
PN J03030678-A.
PD 08-FEB-1991.
PF 29-JUN-1989; 167689.
PR (OHSA-) OHSKA BIOSCIENCE K.
PA WPI; 91-084343/12.
DR N-PSDB; 010957.
PT DNA cfd. coding luciferase derived from Vargula hilgendorffii - by
transfection of host cells and culturing.
PS Claim 1; Fig 3; 11pp; Japanese.
CC The cDNA encoding the protein is used to prepare vector pRSVVL
which is used to transfect COS cells (ATCC CRL1650) for expression
of luciferase. The enzyme is useful for assays in biomedical or
environmental fields.
SQ Sequence 555 AA;

Query Match 18.3%; Score 69; DB 2; Length 555;
Best Local Similarity 33.3%; Pred. No. 2.26e+01;
Matches 8; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

Db 39 ceakegecidrcatckridlsdg 62
QY 13 CSPKOTEGHRCESCMALLSSS 36

Search completed: Tue Mar 19 09:55:19 1996
Job time : 9 secs.

W P E R E H
(TM)

Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by IntelliGenetics, Inc.

MSrch_pp protein - protein database search, using Smith-Waterman algorithm
Run on: Tue Mar 19 09:54:27 1996; MasPar time 3.63 Seconds
Tabular output not generated. 341.191 Million cell updates/sec

Title: >US-08-404-832-3
Description: (1-49) from US08404832.pep
Perfect Score: 377
Sequence: 1 DKYKCECHLVICSPKOTEGHRCESCMALLSSSPKCAOESIVK 49

Scoring table: PAM 150
Gap 11

Searched: 82306 seqs, 25270970 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database:

pir46
1:ann1 2:ann2 3:ann3 4:unann1 5:unann2 6:unann3 7:unann4
8:unann5 9:unann6 10:unann7 11:unann8 12:unrev1
13:unrev2

Statistics: Mean 30.545; Variance 66.324; scale 0.461

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description	Pred. No.
1	377	100.0	568	11	A55649	TNFR-associated prot	8.71e-52
2	120	31.8	377	11	A47380	KING finger-contain	8.45e-07
3	109	28.9	1863	10	A54652	breast/ovarian cance	3.93e-05
4	107	28.4	475	2	A37241	52K autoantigen Ro/S	7.80e-05
5	105	27.9	1042	11	A43379	RAG-1 protein - rabb	1.54e-04
6	105	27.9	1042	11	S42511	RAG-1 protein - rabb	1.54e-04
7	102	27.1	259	9	S53400	18543.3 protein - ye	4.24e-04
8	102	27.1	487	3	DDBY18	DNA repair protein R	4.24e-04
9	102	27.1	1043	11	A33754	recombination-activa	4.24e-04
10	99	26.3	501	12	S34825	uvr2 protein - Neuro	1.16e-03
11	99	26.3	609	10	A43906	nuclear phosphoprote	1.16e-03
12	98	26.0	1043	11	B33754	recombination-activa	1.16e-03
13	97	25.7	477	9	S53494	C3HC4 type zinc fing	2.24e-03
14	94	24.9	328	9	S00530	surface antigen CRP1	5.99e-03
15	94	24.9	498	12	B42125	cysteine-rich surfac	5.99e-03
16	94	24.9	624	10	S28418	probable zinc-bindin	5.99e-03
17	94	24.9	1009	10	A56095	Hip116 protein - hum	5.99e-03
18	94	24.9	1009	10	S49618	helicase-like transc	5.99e-03
19	94	24.9	1041	13	S42509	Rag-1 protein - chic	5.99e-03
20	90	23.9	160	10	S28290	hypothetical protein	2.18e-02
21	85	22.5	658	8	A44272	photomorphogenesis r	1.06e-01
22	85	22.5	834	10	S44866	ROS53.4 protein - Ca	1.06e-01
23	85	22.5	1169	9	S31301	DNA repair protein R	1.06e-01
24	83	22.0	315	6	D48560	immediate-early prot	1.97e-01
25	83	22.0	398	7	S34203	heat shock protein 7	1.97e-01
26	83	22.0	598	7	S28712	heat shock protein 7	1.97e-01
27	83	22.0	630	10	A49656	estrogen-responsive	1.97e-01
28	83	22.0	775	3	EDBE11	immediate-early prot	1.97e-01
29	82	21.8	361	13	S17880	gene posterior sex c	2.68e-01
30	82	21.8	365	11	A30891	regulatory protein r	2.68e-01
31	82	21.8	1603	10	S17963	gene posterior sex c	2.68e-01


```

##molecule_type DNA
##residues 1-1042 ##label FUS
##cross-references EMBL:M7666

SUMMARY #length 1042 #molecular-weight 118992 #checksum 8787

Query Match 27.9%; Score 105; DB 11; Length 1042;
Best Local Similarity 35.0%; Pred. No. 1.54e-04;
Matches 14; Conservative 7; Mismatches 19; Indels 0; Gaps 0;

Db 292 cgitcenhldqvetschvforicilliclkmgyscpqc 331
      1 : : : : : : : : : : : : : : : : : :
Qy 5 CERKHIVLCSPKQTECGHRCFCSCMALSSSPKCTACQ 44

RESULT 7
ENTRY S53400 #type complete
TITLE L8943.3 protein - yeast (Saccharomyces cerevisiae)
ORGANISM "formal name Saccharomyces cerevisiae"
DATE 05-May-1995 #sequence_revision 01-Sep-1995 #text_change
01-Sep-1995
ACCESSIONS S53400
REFERENCE S53390
AUTHORS Du, Z.
SUBMISSION submitted to the EMBL Data Library, February 1995
DESCRIPTION The sequence of S. cerevisiae cosmid 8543.
ACCESSION S53400
STATUS Preliminary
MOLECULE_TYPE DNA
RESIDUES 1-259 ##label DQZ
CROSS-REFERENCES EMBL:D20618

SUMMARY #length 259 #molecular-weight 29741 #checksum 8279

Query Match 27.1%; Score 102; DB 9; Length 259;
Best Local Similarity 40.5%; Pred. No. 4.24e-04;
Matches 17; Conservative 6; Mismatches 18; Indels 1; Gaps 1;

Db 197 fkcitckedykspvncdghfcgsfakdkkgdt-kcfich 237
      : : : : : : : : : : : : : : : : : :
Qy 3 YKCEKHIVLCSPKQTECGHRCFCSCMALSSSPKCTACQ 44

RESULT 8
ENTRY DDBy18 #type complete
TITLE DNA repair protein Rad18 - yeast (Saccharomyces cerevisiae)
INTERMEDIATE_NAMES Protein YCRB66w
ORGANISM "formal name Saccharomyces cerevisiae"
DATE 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change
12-May-1995
ACCESSIONS S05802; S22263; S19481; JS0082
REFERENCE S05802
AUTHORS Jones, J.S.; Weber, S.; Prakash, L.
JOURNAL Nucleic Acids Res. (1988) 16:7119-7131
TITLE The Saccharomyces cerevisiae Rad18 gene encodes a protein
that contains potential zinc finger domains for nucleic
acid binding and a putative nucleotide binding sequence.
CROSS-REFERENCES MIDB:8830333
ACCESSION S05802
MOLECULE_TYPE DNA
RESIDUES 1-487 ##label JUN
CROSS-REFERENCES EMBL:X12588

```

```

REFERENCE      S22260
#authors      Benit, P.; Chanet, R.; Fabre, F.; Faye, G.; Fukuhara, H.;
#journal      Sor, F.
#title        Yeast (1992) 8:147-153
#cross-references EMBL:92221691
#accession    S22263
##molecule_type DNA
##residues    1-487 ##label BEN
#cross-references EMBL:S93798
REFERENCE      S19477
#authors      Antoline, G.; Benit, P.; Chanet, R.; Fabre, R.; Faye, G.;
#submission    Fukuhara, H.; Mathieu, A.; Sor, F.
#accession     Submitted to the Protein Sequence Database, March 1992
#molecule_type DNA
##residues    1-487 ##label ANT
#cross-references EMBL:X59720
REFERENCE      JS0082
#authors      Chanet, R.; Magana-Schwenke, N.; Fabre, F.
#journal       Gene (1988) 74:343-347
#title        Potential DNA-binding domains in the RAD18 gene product of
               Saccharomyces cerevisiae.
#cross-references M01D:89232745
#accession     JS0082
##molecule_type DNA
##residues    1-487 ##label CHA
GENETICS
#gene          LISTA:RAD18
#map_position  3R
CLASSIFICATION #superfamily DNA repair protein RAD18
KEYWORDS       ATP; DNA binding; DNA repair; zinc finger
FEATURE
28-48          #region zinc finger motif\
51-65          #region zinc finger motif\
190-210        #binding site ATP (Lys) #status predicted
366            #length 487 #molecular-weight 55230 #checksum 1135
SUMMARY
Query Match      27.1%; Score 102; DB 3; Length 487;
Best Local Similarity 37.5%; Pred. No. 4.24e-04;
Matches 15; Conservative 6; Mismatches 18; Indels 1; Gaps 1;

Dy      27 rchckdflkpyltpcgtfscilcithlmg-pncplc 65
      :| | | | | | | | | | | | | | | | | | | |
      4 KCCKHIVLCSPKQTCGHRFCESCMALLSSSEPKTAC 43

RESULT      9
ENTRY       A33754      #type complete
TITLE       recombination-activating protein 1 - human
ORGANISM    #formal_name Homo sapiens #common_name man
DATE        07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change
              31-Dec-1993
ACCESSIONS  A33754
REFERENCE    A33754
#authors     Schatz, D.G.; Oettinger, M.A.; Baltimore, D.
#journal     Cell (1989) 59:1035-1046
#title       The V(D)J recombination activating gene, RAG-1.

```


Page
66

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description	Pred. No.
1	176	46.7	501	7	TRF2_MOUSE	TNF RECEPTOR ASSOCIAT	1.71e-19
2	120	31.8	377	6	RINI_HUMAN	BREAST CANCER TYPE 1	1.11e-08
3	109	28.9	1863	6	ROS2_HUMAN	52 KD RO PROTEIN (SIO	1.06e-06
4	107	28.4	475	6	ROS2_HUMAN	V(D)J RECOMBINATION A	5.32e-06
5	105	27.9	1042	6	RAG1_RABIT	DNA REPAIR PROTEIN RA	1.76e-05
6	102	27.1	487	6	RAG1_YEAST	V(D)J RECOMBINATION A	1.76e-05
7	102	27.1	1043	6	RAG1_HUMAN	V(D)J RECOMBINATION A	1.76e-05
8	99	26.3	501	7	UVS2_NEOCR	SURFACE ANTIGEN CRP17	8.52e-05
9	98	26.0	1040	6	RAG1_MOUSE	V(D)J RECOMBINATION A	4.01e-04
10	94	24.9	328	1	C170_GAIA	ZINC-BINDING PROTEIN	4.01e-04
11	94	24.9	624	1	A33_FLEMA	V(D)J RECOMBINATION A	1.84e-03
12	94	24.9	1041	6	RAG1_CHICK	HYPOTHETICAL 18.7 KD	1.84e-03
13	90	23.9	160	8	YNN1_CAEL	HYPOTHETICAL 97.1 KD	1.18e-02
14	85	22.5	834	6	YNN4_CAEL	DNA REPAIR PROTEIN RA	2.43e-02
15	83	22.5	1169	6	RAD5_YEAST	HEAT SHOCK PROTEIN 70	2.43e-02
16	83	22.0	598	4	HS7L_SHRV	TRANS-ACTING TRANSCRI	3.49e-02
17	82	22.0	775	4	ICP0_HSV1	DOWN REGULATORY PROTE	3.49e-02
18	82	21.8	353	6	RPT1_MOUSE	POSTERIOR SEX COMBS P	3.49e-02
19	82	21.8	1603	6	PSC_DROME	MAJOR SURFACE-LABELLED	5.00e-02
20	81	21.5	713	7	TS4A_GAIA	CYSTEINE-RICH PROTEIN	1.02e-01
21	79	21.0	208	2	CRP2_RAT	DP87 PROTEIN	1.02e-01
22	79	21.0	555	2	DP87_DICDI	INTEGRIN BETA-1 SUBUN	1.44e-01
23	78	20.7	798	4	ITB1_XENLA	INTEGRIN BETA-1* SUBU	1.44e-01
24	78	20.7	798	4	ITB1_XENLA	DNA-BINDING PROTEIN B	2.03e-01
25	77	20.4	324	1	BM1_MOUSE	HYPOTHETICAL 30.8 KD	2.89e-01
26	77	20.4	326	1	BM1_MOUSE	PUTATIVE TRANSCRIPTIO	2.89e-01
27	76	20.2	274	8	YEO6_YEAST	PROBABLE TRANSCRIPTIO	2.89e-01
28	76	20.2	560	5	PML2_HUMAN	PROBABLE TRANSCRIPTIO	2.89e-01
29	76	20.2	589	8	Y041_CAEL	PROBABLE TRANSCRIPTIO	2.89e-01
30	76	20.2	611	5	PML2_HUMAN	PROBABLE TRANSCRIPTIO	2.89e-01
31	76	20.2	633	5	PML3_HUMAN	PROBABLE TRANSCRIPTIO	2.89e-01
32	76	20.2	824	5	PML3_HUMAN	PROBABLE TRANSCRIPTIO	2.89e-01
33	76	20.2	882	5	PML3_HUMAN	PROBABLE TRANSCRIPTIO	2.89e-01
34	76	20.2	1429	4	LI12_CAEL	LIN-12 PROTEIN PRECUR	4.08e-01
35	75	19.9	115	8	YBR2_YEAST	HYPOTHETICAL 13.2 KD	4.08e-01
36	75	19.9	513	6	PSV_DICDI	PRESPORE VESICLE PROT	4.08e-01
37	74	19.6	803	4	ITB1_CHICK	INTEGRIN BETA-1 PRECU	5.73e-01
38	74	19.6	1801	4	IMB2_RAT	LAMININ BETA-2 CHAIN	8.03e-01
39	73	19.4	138	5	PA2B_VIPAA	PHOSPHOLIPASE A2 PREC	8.03e-01
40	73	19.4	138	5	PA2C_VIPAA	PHOSPHOLIPASE A2 PREC	8.03e-01
41	73	19.4	138	5	PA2A_VIPAA	PHOSPHOLIPASE A2 PREC	8.03e-01
42	73	19.4	236	3	G1EM_TYLAH	GIANT EXTRACELLULAR H	8.03e-01
43	73	19.4	342	4	ME18_MOUSE	DNA-BINDING PROTEIN M	8.03e-01
44	73	19.4	344	4	ME18_MOUSE	DNA-BINDING PROTEIN M	8.03e-01
45	73	19.4	410	4	ICP0_PRIVIF	TRANS-ACTING TRANSCRI	8.03e-01

ALIGNMENTS

RESULT 1
ID TRF2_MOUSE STANDARD; PRT; 501 AA.
AC P39429;
DT 01-FEB-1995 (REL. 31, CREATED)

DT	01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)
DT	01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)
DE	TNF RECEPTOR ASSOCIATED FACTOR 2 (TRAF2).
OS	MUS MUSCULUS (MOUSE).
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
CC	EUTHERIA; RODENTIA.
CC	[1]
RP	SEQUENCE FROM N.A.
RM	94349371
RA	ROTHE M., WONG S.C., HENZEL W.J., GOEDEL D.V.;
RL	CELL 78:681-692(1994).
CC	-1- FUNCTION: SIGNAL TRANSDUCER ASSOCIATED WITH THE CYTOPLASMIC DOMAIN
CC	OF THE 75 KD TUMOR NECROSIS FACTOR RECEPTOR (TNF-R2).
CC	-1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC	-1- SUBUNIT: HETERODIMER OF TRAF1 AND TRAF2.
CC	-1- SIMILARITY: CONTAINS A C3HC4-CLASS ZINC FINGER.
DR	EMBL; L33303; MMTAF2A.
KW	ZINC-FINGER; COILED COIL.
FT	ZN FING 34
SQ	SEQUENCE 501 AA; 56026 MW; 1264823 CN;
Db	31 kyleacknllrrfgagqghrycsfclstslssgpncaac 72
Qy	2 KYRCKCHLVLCSPKQTECHRCFCSCMALLSSSPKCTAC 43
Query Match	46.7%; Score 176; DB 7; Length 501;
Best Local Similarity	45.2%; Pred. No. 1,71e-19;
Matches	19; Conservative 12; Mismatches 11; Indels 0; Gaps 0;
RESULT 2	
ID	RINI_HUMAN STANDARD; PRT; 377 AA.
AC	Q06587;
DT	01-JUN-1994 (REL. 29, CREATED)
DT	01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)
DT	01-OCT-1994 (REL. 30, LAST ANNOTATION UPDATE)
DE	RING1 PROTEIN.
GN	RING1.
OS	HOMO SAPIENS (HUMAN).
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
CC	EUTHERIA; PRIMATES.
CC	[1]
RP	SEQUENCE FROM N.A.
RM	93211912
RA	LOVERING R., HANSON I.M., BORDEN K.L.B., MARTIN S., O'REILLY N.J.;
RA	EVAN G.I., RAHMAN D., PAPPIN D.J.C., THOMSDALE J., FREMONT P.S.;
RL	PROC. NATL. ACAD. SCI. U.S.A. 90:2112-2116(1993).
CC	-1- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).
CC	-1- SIMILARITY: CONTAINS A C3HC4-CLASS ZINC FINGER.
DR	EMBL; Z14000; HSRING1.
DR	PIR; A47380; A47380.
DR	HSSP; P28990; ICHC.
KW	PROSITE; PS00518; ZINC FINGER C3HC4.
KW	ZINC-FINGER; DNA-BINDING; NUCLEAR PROTEIN.
FT	ZN FING 19
FT	DOMAIN 176 231
FT	DOMAIN 285 348
FT	DOMAIN 172 175
SQ	SEQUENCE 377 AA; 39145 MW; 716288 CN;

CC	-1-	DISEASE: SERA FROM PATIENTS WITH SYSTEMIC LUPUS ERYTHEMATOSUS OR PRIMARY SJOGEN'S SYNDROME OFTEN CONTAIN ANTIBODIES THAT REACT WITH NORMAL CELLULAR RO PROTEIN AS IF THIS ANTIGEN WAS FOREIGN.
CC	-1-	SUBCELLULAR LOCATION: CYTOPLASMIC AND NUCLEAR.
CC	-1-	SIMILARITY: CONTAINS A C3HC4-CLASS ZINC FINGER.
DR	EMBL:	M34551; HSRSSA.
DR	EMBL:	M62800; HSSSARO.
DR	PIR:	A37241; A37241.
DR	MM:	109092; 11TH EDITION.
DR	PROSITE:	PS00518; ZINC FINGER C3HC4.
KW	RIBONUCLEOPROTEIN;	ZINC-FINGER; ANTIGEN; RNA-BINDING;
FT	DOMAIN	1 24 HYDROPHOBIC.
FT	ZN FING	16 54 C3HC4-TYPE.
FT	DOMAIN	92 123 B BOX.
FT	DOMAIN	211 232 LEDICINE-ZIPPER.
FT	DOMAIN	401 475 HYDROPHOBIC.
FT	CONFLICT	52 52 P -> A (IN REF. 2).
SO	SEQUENCE	475 AA; 54169 MW; 1162917 CN;
Db	16	cpcldpfvpeyviecghafceciagvgkgygsevcprctqrlf 59
Oy	5	CERCHLVLCSPKOTECGHRFCSCMAALLSSSPKTCACQESIV 48
RESULT	5	
ID	RAG1 RABIT	STANDARD; ERT; 1042 AA.
AC	P34088;	
DT	01-FEB-1994	(REL. 28, CREATED)
DT	01-FEB-1994	(REL. 28, LAST SEQUENCE UPDATE)
DT	01-FEB-1994	(REL. 28, LAST ANNOTATION UPDATE)
DE	V(D)J RECOMBINATION ACTIVATING PROTEIN.	
GN	RAG-1.	
OS	ORYZOLUGUS CONICULUS (RABBIT).	
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;	
OC	ETHERIA; LAGOMORPHA.	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RM	93354283	
RA	FISCHIOITI P., HARINDRANATH N., MAGE R.G., MCCORMACK W.T.,	
RL	DHANAPALAN P., RODX K.H.;	
RU	MOL. IMMUNOL. 30:1021-1032(1993).	
CC	-1-	FUNCTION: RAG1 & RAG2 SYNERGISTICALLY ACTIVATE THE IMMUNOGLOBULIN V-D-J RECOMBINATION. V-D-J RECOMBINATION IS THE COMBINATORIAL PROCESS BY WHICH DEVELOPING LYMPHOCYTES BEGIN TO GENERATE THEIR ENORMOUS RANGE OF BINDING SPECIFICITIES FROM A LIMITED AMOUNT OF GENOMIC INFORMATION.
CC	-1-	SUBCELLULAR LOCATION: NUCLEAR.
CC	-1-	SIMILARITY: CONTAINS A C3HC4-CLASS ZINC FINGER.
DR	EMBL:	M77666; OCRAGIO.
DR	PIR:	S44379; S44379.
DR	PROSITE:	PS00518; ZINC FINGER C3HC4.
KW	ZINC-FINGER; DNA-BINDING; NUCLEAR PROTEIN.	
FT	ZN FING	292 330 C3HC4-TYPE.
SO	SEQUENCE	1042 AA; 119006 MW; 5538604 CN;

Query Match	Similarity	Score	DB	Length
Best Local	14%	Conservative	7%	Mismatches 19%
Matches	14%	Conservative	7%	Mismatches 19%
Indels	0%	Gaps	0%	
Db	292	cgicenhlladvetsckhvfcrclrlclckwmsgycpsqc	331	
Qy	5	CEKCHLVLCSPKTEGCHRFCESCMAALLSSSPKCTACQ	44	
RESULT	6	STANDARD;	PRT;	487 AA.
ID	RA18	YEAST		
AC	P10862;			
DT	01-JUL-1989	(REL. 11, CREATED)		
DT	01-JUL-1989	(REL. 11, LAST SEQUENCE UPDATE)		
DT	01-FEB-1994	(REL. 28, LAST ANNOTATION UPDATE)		
DE	DNA REPAIR PROTEIN RAD18.			
DE	RAD18 OR YCR66W.			
OS	SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).			
OC	EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOCETES.			
LN	[1]			
RE	SEQUENCE FROM N.A.			
RM	92221691			
RA	BENIT P., CHANET R., FABRE F., FAYE G., FUKUHARA H., SOR F.,			
RL	YEAST 8:147-153(1992).			
LN	[2]			
RE	SEQUENCE FROM N.A.			
RM	89232745			
RA	CHANET R., MAGANA-SCHWENCKE N., FABRE F.,			
RL	GENE 74:543-547(1988).			
LN	[3]			
RE	SEQUENCE FROM N.A.			
RM	88303333			
RA	JONES J.S., WEBER S., PRAKASH L.;			
RL	NUCLEIC ACIDS RES. 16:7119-7131(1988).			
CC	-1- FUNCTION: FUNCTIONS WITH DNA REPAIR PROTEIN RAD5 IN ERROR-FREE			
CC	POSTREPLICATION DNA REPAIR. RAD5 INCREASES EFFICIENCY OF RAD18			
CC	FUNCTION.			
CC	-1- SIMILARITY: CONTAINS A C3HC4-CLASS ZINC FINGER.			
CC	-1- SUBCELLULAR LOCATION: NUCLEAR.			
DR	EMBL; X12586; SCRAD18.			
DR	EMBL; X59720; SCRH11.			
DR	PIR; S05802; DDBY18.			
DR	PIR; J50082; J50082.			
DR	LISTA; SC00917; RAD18.			
DR	PROSITE; PS00518; ZINC FINGER C3HC4.			
KW	ZINC-FINGER; DNA-BINDING; DNA DAMAGE; DNA REPAIR; NUCLEAR PROTEIN.			
FT	ZN FING 28 65 C3HC4-TYPE.			
FT	ZN FING 190 210 C2HC-TYPE.			
SQ	SEQUENCE 487 AA; 55230 MW; 1203687 CN;			
Query Match	27.1%;	Score 102;	DB 6;	Length 487;
Best Local	Similarity 37.5%;	Pred. No. 1.76e-05;		
Matches	15;	Conservative 6;	Mismatches 18;	Indels 1;
Gaps	1;			
Db	27	rchickdflkvpvlpcqgntfscloirhlnq-pncpic	65	
Qy	4	KCEKCHLVLCSPKTEGCHRFCESCMAALLSSSPKCTAC	43	

ID	RAG1 HUMAN	STANDARD;	PRT;	1043 AA.
AC	P15918;			
DT	01-APR-1990 (REL. 14, CREATED)			
DT	01-APR-1990 (REL. 14, LAST SEQUENCE UPDATE)			
DT	01-NOV-1991 (REL. 20, LAST ANNOTATION UPDATE)			
DE	V(D)-J RECOMBINATION ACTIVATING PROTEIN.			
GN	RAG1.			
OS	HOMO SAPIENS (HUMAN).			
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA; EUTHERIA; PRIMATES.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RM	9009604			
RA	SCHATZ D.G., OETTINGER M.A., BALTIMORE D.;			
RL	CELL 59:1035-1046(1989).			
CC	-I- FUNCTION: RAG1 & RAG2 SYNERGISTICALLY ACTIVATE THE IMMUNOGLOBULIN V-D-J RECOMBINATION. V-D-J RECOMBINATION IS THE COMBINATORIAL PROCESS BY WHICH DEVELOPING LYMPHOCYTES BEGIN TO GENERATE THEIR ENDLESS RANGE OF BINDING SPECIFICITIES FROM A LIMITED AMOUNT OF GENETIC INFORMATION.			
CC	-I- SUBCELLULAR LOCATION: NUCLEAR.			
CC	-I- TISSUE SPECIFICITY: MATURING LYMPOID CELLS.			
DR	-I- SIMILARITY: CONTAINS A C3HC4-CLASS ZINC FINGER.			
DR	EMBL, M29474; HSRAG1.			
DR	PIR; A33754; A33754.			
DR	HSPR; P28990; 1CHC.			
DR	MIM; 119615; 11TH EDITION.			
DR	PROSITE; PS00518; ZINC_FINGER_C3HC4.			
KW	ZINC-FINGER; DNA-BINDING; NUCLEIC ACID-BINDING PROTEIN.			
FT	ZN FING 293 331 C3HC4-TYPE.			
SO	SEQUENCE 1043 AA; 119115 MW; 5512008 CN;			
Query Match	27.1%;	Score 102;	DB 6;	Length 1043;
Best Local Similarity	32.5%;	Pred.No. 1.76e-05;		
Matches 13;	Conservative 9;	Mismatches 18;	Indels 0;	Gaps 0;
Dn	293 cdcicenh1adppvncnkhyfcrvcilrcikymgdycpscr 332	: : : : : : : :		
Oy	5 CERKHYLCSPKOTEGCRHFCSCMAALLSSSPSKTCAQ 44	: : : : : : : :		
RESULT	8			
ID	IWS2 NEUCR	STANDARD;	PRT;	501 AA.
AC	P33288;			
DT	01-FEB-1994 (REL. 28, CREATED)			
DT	01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)			
DT	01-FEB-1994 (REL. 28, LAST ANNOTATION UPDATE)			
DE	IWS-2 PROTEIN.			
GN	IWS-2.			
OS	NEUROPODA CRASSA.			
OC	EUKARYOTA; FUNGI; ASCOMYCOTINA; PYRENUMYCETES; SORDARIALES.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=C1-110--34A;			
RM	93241158			
RA	TOMITA H., SOSHI T., INOUE H.;			
RL	MOL. GEN. GENET. 238:225-233(1993).			
CC	-I- FUNCTION: MAY BE AN ACTIVATOR THAT FUNCTIONS IN THE TRANSCRIPTION OF DNA REPAIR GENES.			
CC	-I- SUBCELLULAR LOCATION: NUCLEAR (PROBABLE).			

[illegible]

QY 5 CEKCHLVLCSPKOTEGHRCFCSMAALLSSSPKCTACQ 44

RESULT 10
ID C170 G1A1A STANDARD; PRT; 328 AA.
AC P15799;
DT 01-APR-1990 (REL. 14, CREATED)
DT 01-APR-1990 (REL. 14, LAST SEQUENCE UPDATE)
DE SURFACE ANTIGEN CRP170 (FRAGMENT).
OS GIARDIA LAMBLIA.
OC EUKARYOTA; PROTOZOA; SARCOMASTIGOPHORA; MASTIGOPHORA; DIPLOMONADIDA;
OC HEXAMITIDE.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ISOLATE WB;
RM 88089405
RA ADAM R.D., AGGARWAL A., TAL A.A., DE LA CRUZ V.F., MCCUTCHAN T.,
RA NASH T.E.;
RL J. EXP. MED. 167:109-118(1988).
CC -1- CISTEINE-RICH, ANTIGENICALLY VARIANT SURFACE PROTEIN.
DR EMBL; X06741; GICRP170.
DR PIR; S00530; S00530.
DR HSSP; P04355; AMT2.
KM REPEAT; ANTIGEN.
FT NON TER 1 1
FT REPEAT 38 102
FT REPEAT 103 167
FT NON TER 328 328
SQ SEQUENCE 328 AA; 33438 MW; 499276 CN;

Query Match 24.9%; Score 94; DB 1; Length 328;
Best Local Similarity 29.4%; Pred. No. 4.01e-04;
Matches 10; Conservative 10; Mismatches 13; Indels 1; Gaps 1;

Db 211 cpaatpgc-bascdgctenmtgacdtctckeg 243
QY 13 CSPKOTEGHRCFCSMAALLSSSPKCTACQES 46

RESULT 11
ID A33 PLEMA STANDARD; PRT; 624 AA.
AC 002084;
DT 01-OCT-1994 (REL. 30, CREATED)
DT 01-OCT-1994 (REL. 30, LAST SEQUENCE UPDATE)
DT 01-OCT-1994 (REL. 30, LAST ANNOTATION UPDATE)
DE ZINC-BINDING PROTEIN A33.
OS PLEURODELES MALITII (IBERIAN RIBBED NEWT).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; AMPHIBIA;
OC CAUDATA.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=OVARY;
RM 93154311
RA BELLINI M., LACROIX J.-C., GALL J.G.;
RL EMBL J. 12:107-114(1993).
CC -1- FUNCTION: MAY BE A NUCLEAR REGULATORY PROTEIN THAT IS STORED
CC IN THE GERMINAL VESICLE FOR USE DURING EARLY EMBRYOGENESIS
CC AND MAY PLAY A ROLE IN THE SYNTHESIS OR PROCESSING OF PRE-MRNA
CC DURING OOGENESIS.

CC -1- DEVELOPMENTAL STAGE: IT FIRST APPEARS ON THE CHROMOSOME LOOPS
CC AND IN THE NUCLEOLAR PLASMA OF THE GERMINAL VESICLE (GV). IT IS
CC TRANSMITTED TO THE EGG AT GV BREAKDOWN AND APPEARS IN EMBRYONIC
CC NUCLEI AT THE MID-BLASTULA STAGE AND IS FOUND IN MANY BUT NOT
CC ALL NUCLEI AT STILL LATER STAGES OF EMBRYOGENESIS.
CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC -1- SIMILARITY: CONTAINS A C3HC4-CLASS ZINC FINGER.
DR EMBL; L04190; PWPMA33A.
KW ZINC-FINGER; NUCLEAR PROTEIN; DEVELOPMENTAL PROTEIN; COILED COIL;
HEPARD REPEAT PATTERN; RNA-BINDING.
FT DOMAIN 19 26
FT DOMAIN 134 149 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT ZN FING 162 201 C3HC4-TYPE.
FT DOMAIN 238 269 B BOX.
FT DOMAIN 337 386 COILED COIL.
FT DOMAIN 449 619 RFP-LIKE.
SQ SEQUENCE 624 AA; 71056 MW; 2023970 CN;

Query Match 24.9%; Score 94; DB 1; Length 624;
Best Local Similarity 32.6%; Pred. No. 4.01e-04;
Matches 15; Conservative 7; Mismatches 23; Indels 1; Gaps 1;

Db 158 edltcplrtsfkcpvllccgntckhcdkswnesataspcke 203
QY 1 DKYKCEKCHLVLCSPKOTEGHRCFCSMAALLSSSP-KCTACQ 45

RESULT 12
ID P4271; STANDARD; PRT; 1041 AA.
AC P24271;
DT 01-MAR-1992 (REL. 21, CREATED)
DT 01-MAR-1992 (REL. 21, LAST SEQUENCE UPDATE)
DT 01-MAR-1992 (REL. 21, LAST ANNOTATION UPDATE)
DE V(D)J RECOMBINATION ACTIVATING PROTEIN.
GN PAC-1.
OS GALUS GALLUS (CHICKEN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; AVES; NEOGNATHAE;
OC GALLIFORMES.
RN [1]
RP SEQUENCE FROM N.A.
RM 91098648
RA CARLSON L.M., OETTINGER M.A., SCHATZ D.G., MASTELLER E.L.,
RA HURLEY E.A., MCCORMACK W.I., BALTMORE D., THOMPSON C.B.;
RL CELL 64:201-208(1991).
CC -1- FUNCTION: PAC1 & BAG2 SYNERGISTICALLY ACTIVATE THE IMMUNOGLOBULIN
CC V-D-J RECOMBINATION. V-D-J RECOMBINATION IS THE COMBINATORIAL
CC PROCESS BY WHICH DEVELOPING LYMPHOCYTES BEGIN TO GENERATE THEIR
CC ENORMOUS RANGE OF BINDING SPECIFICITIES FROM A LIMITED AMOUNT OF
CC GENETIC INFORMATION.
CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC -1- SIMILARITY: CONTAINS A C3HC4-CLASS ZINC FINGER.
DR EMBL; M58530; GGRAG1G.
DR PIR; S42509; S42509.
DR PROSITE; PS00518; ZINC FINGER C3HC4.
KW ZINC-FINGER; DNA-BINDING; NUCLEAR PROTEIN.
FT ZN FING 284 327 C3HC4-TYPE.
SQ SEQUENCE 1041 AA; 119916 MW; 5442207 CN;

Query Match 24.9%; Score 94; DB 6; Length 1041;
Best Local Similarity 30.8%; Pred. No. 4.01e-04;

CC MAINTENANCE OF WILD-TYPE RATES OF INSTABILITY OF SIMPLE
CC REPETITIVE SEQUENCES SUCH AS POLY(GT) REPEATS.
CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC -1- SIMILARITY: CONTAINS A C3HC4-CLASS ZINC FINGER.
CC -1- SIMILARITY: TO HELICASES OF THE SNF2/RAD54 FAMILY.
DR EMBL; M96644; SCRAD5A.
DR EMBL; S43248; S43248.
DR EMBL; S46103; S46103.
DR PIR; S31301; S31301.
DR PIR; S26983; S26983.
DR LISTA; SC00924; RAD5.
DR PROSITE; PS00516; ZINC_FINGER_C3HC4.
KW DNA DAMAGE; DNA REPAIR; NUCLEAR PROTEIN; ZINC-FINGER; DNA-BINDING;
KW HELICASE; ATP-BINDING.
FT DOMAIN 42 60 ASP/GLU-RICH (ACIDIC).
FT NP BIND 303 315 ARG/LYS-RICH (BASIC).
FT SITE 532 539 ATP (POTENTIAL).
FT ZN FING 681 684 DECH BOX.
FT ZN FING 914 960 C3HC4-TYPE.
FT CONFLICT 478 478 Q -> R (IN REF. 2).
FT CONFLICT 635 635 T -> N (IN REF. 2).
FT CONFLICT 846 846 G -> S (IN REF. 2).
FT CONFLICT 898 898 R -> S (IN REF. 2).
FT CONFLICT 973 973 V -> A (IN REF. 2).
FT CONFLICT 1063 1063 A -> R (IN REF. 2).
SQ SEQUENCE 1169 AA; 134001 MW; 7092987 CN;

Query Match 22.5%; Score 85; DB 6; Length 1169;
Best Local Similarity 75.0%; Pred. No. 1.18e-02;
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Db 930 tecghsfecsc1 941
||| ||| ||| |
QY 18 TECGHRFECSCM 29

Search completed: Tue Mar 19 09:54:09 1996
Job time : 7 secs.